

The survival of organisms depends on their ability to perceive and respond to extra-cellular signals. At the molecular level, signals are perceived and transmitted through networks of interacting proteins or the like, that act cooperatively to maintain cellular homeostasis and regulate activities like growth, division and differentiation. Information transfer through biological signaling networks is mediated largely by protein-protein interactions that can assemble and disassemble dynamically in response to signals, creating transient circuits that link external events to specific internal outputs, such as changes in gene expression. Numerous strategies have been developed to map the

protein-protein interactions that underlie these networks. These studies have collectively provided a wealth of data delineating genome-wide protein-protein interactions for *Saccharomyces cerevisiae* and other organisms. While  
5 powerful, these approaches have provided only a partial picture and are likely to overlook many interactions that are context dependent, forming only in the presence of their appropriate signals.

10                   The disruption of protein-protein interactions either by mutation or small-molecules can create biological fulcrums that enable small perturbations of a signaling network to elicit large changes in cellular phenotype. However, not all protein-protein interactions in a given  
15 signaling pathway are likely to possess this power. As such, complementary strategies that aim to identify regulatory protein-protein interactions by artificially introducing proteins or peptides into cells which compete with and titrate-out the endogenous regulatory interactions,  
20 thereby disrupting the normal circuits that connect external signals to cellular responses, are of interest. By combining this strategy with functional assays, such as the activation of a gene in response to a signal, screens for functional interference can be used to identify peptides  
25 that perturb regulatory protein-protein interactions. This strategy, often referred to as dominant-interfering or dominant-negative genetics, has been successfully employed in several model organisms where high-throughput screening methods are easily applied, and to a lesser extent  
30 in mammals, which have traditionally been less amenable to

these types of screens. One advantage of dominant-negative strategies is that such strategies can pinpoint the functionally relevant protein-protein interaction "fulcrum points" and thereby expose the small number of nodes within  
5 the larger web of a protein network that are susceptible to functional modulation by external agents. As such, the results of such strategies can provide vital information about the regulatory components that define a particular pathway and can allow the elucidation of key protein-protein  
10 interactions suitable for targeting by drug screening programs.

The difficulty in transfecting cells or producing transgenic organisms hinders the progression of  
15 development of dominant negative screening in mammals. To overcome this problem, high-efficiency retrovirus transfection has been developed. Although this retrovirus transfection is potent, it is necessary to produce DNA to be packaged into viral intermediates, and therefore, the  
20 applicability of this technique is limited. Alternatively, high-density transfection arrays or cell arrays are being developed, and the use thereof is proposed.

Rosetta Inpharmatics has proposed using  
25 cellular information as a profile in some patent applications (Japanese PCT National Phase Laid-Open Publication No.: 2003-505038 ; Japanese PCT National Phase Laid-Open Publication No.: 2003 - 505022 ; Japanese PCT National Phase Laid-Open Publication No.: 2002-533701 ;  
30 Japanese PCT National Phase Laid-Open Publication No.: 2002

—533700 ; Japanese PCT National Phase Laid-Open Publication  
No.: 2002—533699 ; Japanese PCT National Phase Laid-Open  
Publication No.: 2002—528095 ; Japanese PCT National Phase  
Laid-Open Publication No.: 2002 — 526757 ; Japanese PCT  
5 National Phase Laid-Open Publication No.: 2002—518021 ;  
Japanese PCT National Phase Laid-Open Publication No.: 2002  
—518003 ; Japanese PCT National Phase Laid-Open Publication  
No.: 2002—514804 ; Japanese PCT National Phase Laid-Open  
Publication No.: 2002—514773 ; Japanese PCT National Phase  
10 Laid-Open Publication No.: 2002—514437). In such a profile,  
information from separate cells is processed as a group of  
separate pieces of information, but not continuous  
information. Therefore, this technique is limited in that  
information analysis is not conducted on a single (the same)  
15 cell. Particularly, in this technique, analysis is  
conducted only at one specific time point before and after  
a certain change, and a series of temporal changes in a point  
(gene) are not analyzed.

20               Recent advances in profiling techniques have  
led to accurate measurement of cellular components, and thus,  
profiling of cellular information (e.g., Schena et al., 1995,  
"Quantitative monitoring of gene expression patterns with  
a complementary DNA microarray", Science 270:467-470;  
25 Lockhart et al., 1996, "Expression monitoring by  
hybridization to high-density oligonucleotide arrays",  
Nature Biotechnology 14:1675-1680; Blanchard et al., 1996,  
"Sequence to array: Probing the genome's secrets", Nature  
Biotechnology 14:1649; and US 5,569,588). For organisms  
30 whose genome is entirely known, it is possible to analyze



the transcripts of all genes in a cell. In the case of other organisms, for which the amount of known genomic information is increasing, a number of genes in a cell can be simultaneously monitored.

5

As array technology advances, arrays also have been utilized in the field of drug screening (e.g., Marton et al., "Drug target validation and identification of secondary drug target effects using Microarrays", Nat. Med., 1998 Nov, 4(11):1293-301; and Gray et al., 1998, "Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors", Science, 281:533-538). Analysis using profiles (e.g., US Patent No. 5,777,888) and clustering of profiles provides information about conditions of cells, transplantation, target molecules and drug candidates, and/or the relevant functions, efficacy and toxicity of drugs. These techniques can be used to determine a common profile which represents ideal drug activity and disease conditions. Comparing profiles assists in detecting diseases in patients at an early stage, and provides prediction of improved clinical outcomes for patients who have been diagnosed as having a disease.

However, to date, there has been no technique which can provide information about the same cell in the true sense. In the above-described techniques, data is obtained as average for a group of heterologous cells. Analyses and evaluations based on such data lack accuracy. Therefore, there is an increasing demand for a method of providing information at the cellular level.

An object of the present invention is to provide a method for obtaining information, profiles or data of a cell. Another object of the present invention is to provide a method for obtaining information and data relating to cell status in a consistent environment, and a method and system for accurately presenting such data. In particular, a particular purpose of the present invention is to provide a system and a method for directly or as such information of a cell in a consistent environment in terms of complex system information, and providing such data and data sequencing technology *per se*. Another object of the present invention is to provide a digital cell and uses thereof.

#### SUMMARY OF THE INVENTION

The above-mentioned objects have been achieved or solved by immobilizing a cell onto a substrate, monitoring a biological agent, or a collection thereof, on or in a cell over time to produce cell profile data. As such, sequential collection of data from a cell has been enabled, and production of a digital cell has been achieved.

The above-mentioned object has further been achieved or solved by providing a substrate locating a plurality of cells in a consistent environment. Such a substrate has been provided by immobilizing a cell onto a substrate by means of for example, a salt or an actin-like substance, preferably both a salt and an actin-like substance. As such, the present invention allows one to collect cell profile

data from cells of the same type, located in a consistent environment, simultaneously under the same conditions.

Accordingly, the present invention provides the  
5 following:

1. A method for producing profile data relating to information of a cell in a consistent environment, said method comprising the steps of:

10 a) locating a plurality of cells to a support which is capable of maintaining the cells in a consistent environment; and b) monitoring a biological agent, or a collection thereof, on or in the cell to produce the profile data for the cell.

15

2. A method according to Item 1, wherein the biological agent is a nucleic acid molecule or a molecule derived from the nucleic acid molecule.

20 3. A method according to Item 1, wherein the cell is immobilized to the support by a composition comprising a) a complex with a positively charged substance and a negatively charged substance; and b) a salt.

25 4. A method according to Item 1 wherein the cell is provided with an actin-like substance.

5. A method according to Item 1, wherein the cell is immobilized to the support by a composition comprising a)  
30 a complex with a positively charged substance and a

negatively charged substance; and b) a salt, and is provided with an actin-like substance.

6. A method according to Item 1 wherein the biological  
5 agent is selected from the group consisting of a nucleic acid molecule, a protein, a saccharide, a lipid, a low molecular weight molecule, and a complex thereof.

7. A method according to Item 1, wherein the cell is  
10 cultured at least about three days before the step of monitoring.

8. A method according to Item 1, wherein the biological  
agent comprises a nucleic acid molecule encoding a gene.  
15

9. A method according to Item 1, wherein the profile  
comprises a profile for gene expression.

10. A method according to Item 1, wherein the profile  
20 comprises a profile of an apoptosis signal.

11. A method according to Item 1 wherein the profile is  
a profile of a stress signal.

25 12. A method according to Item 1 wherein the profile is  
a profile of the localization of a molecule.

13. A method according to Item 12 wherein the molecule is  
detected by means selected from the group consisting of  
30 fluorescence, phosphorescence, radioactivity, and a

combination thereof.

14. A method according to Item 1 wherein the profile comprises a variation in cell morphology.

5

15. A method according to Item 1 wherein the profile comprises a profile of promoter activity.

16. A method according to Item 1, wherein said profile  
10 comprises a profile of a promoter dependent on a specific drug.

17. A method according to Item 1 wherein said profile  
15 comprises a profile of a promoter dependent on a specific drug, wherein said method further comprises the step of administering the specific drug.

18. A method according to Item 1 further comprising the  
20 step of subjecting the cell to a foreign agent.

19. A method according to Item 18, wherein said foreign agent comprises an RNAi.

20. A method according to Item 18, wherein said foreign  
25 agent comprises a chemical not present in a biological body.

21. A method according to Item 1, wherein said profile comprises a profile of an intermolecular interaction.

30 22. A method according to Item 18, wherein said foreign

agent comprises a ligand for a receptor of said cell.

23. A method according to Item 1, wherein said profile  
comprises a profile of an interaction between a receptor  
5 and a ligand.

24. A method according to Item 1, wherein said profile is  
of a cellular form, and said method further comprises the  
step of giving to said cell a stimulus selected from the  
10 group consisting of overexpression, underexpression or  
knockdown of a gene, addition of a foreign agent and a change  
in the environment.

25. A method according to Item 1, wherein said profile  
15 comprises a profile of interaction between molecules present  
in said cell.

26. A method according to Item 1, further comprising the  
step of conducting observation using a technology selected  
20 from the group consisting of two-hybrid method, FRET and  
BRET.

27. A method according to Item 1, wherein said profile  
comprises a profile of interaction between molecules present  
25 in said cell, wherein the method further comprises the step  
of conducting observation using a technology selected from  
the group consisting of two-hybrid method, FRET and BRET.

28. A method according to Item 1, wherein said cell is  
30 located on said support in an array format.

29. A method according to Item 1, wherein said cell is located on said support in an array format, and each of said plurality of cells are located at a space of 1 mm at maximum.

5

30. A method according to Item 1, wherein said profile is obtained in real time.

31. A method according to Item 1 further comprising the step of immobilizing said cell to a solid support.

10

32. A method according to Item 1, wherein said data comprises information relating to said profile.

33. A method according to Item 1 wherein said data comprises information relating to conditions during said monitoring.

15

34. A method according Item 1 wherein said data comprises information relating to the status of said cell.

20

35. A method according to Item 1 wherein said biological agent to be monitored comprises at least two types of biological agent.

25

36. A method according to Item 1 wherein said biological agent to be monitored comprises at least three types of biological agent.

37. A method according to Item 1 wherein said biological

30

agent to be monitored comprises at least eight types of biological agents.

38. A method according to Item 1 further comprising the  
5 step of arbitrarily selecting a biological agent.

39. A method according to Item 1 wherein said cell is  
selected from the group consisting of a stem cell and a  
somatic cell.  
10

40. A method according to Item 1 wherein said support  
comprises a solid support.

41. A method according to Item 1 wherein said support  
15 comprises a substrate.

42. A method according to Item 1 wherein said biological  
agent is a nucleic acid molecule, and said cell is  
transfected with said nucleic acid molecule.  
20

43. A method according to Item 42, wherein said  
transfection is conducted on a solid phase or in a liquid  
phase.

25 44. A method according to Item 42, wherein said  
transfection is conducted on a solid support.

45. A method according to Item 1, further comprising the  
step of comparing a phase of said profile.  
30



46. A method according to Item 1, further comprising the step of subtracting a control profile from the profile of said cell.

5 47. A method according to Item 1 further comprising the step of processing the profile with a mathematical processing method selected from signal processing and multivariant analysis methods.

10 48. A method for presenting profile data relating to information of a cell in a consistent environment, comprising the steps of:

a) locating a plurality of cells on a support which is capable of maintaining the cells in a consistent  
15 environment;

b) monitoring a biological agent, or a collection thereof, on or in the cell to produce the profile data for the cell; and

c) presenting the data.  
20

49. A method according to Item 48, wherein said step of presenting is conducted in real-time.

50. A method according to Item 48, wherein said step of  
25 presenting is conducted such that visual detection is enabled.

51. A method according to Item 48, wherein said step of  
30 presenting is conducted such that auditory detection is enabled.

52. A method for determining the state of a cell in a consistent environment, comprising the steps of:

5 a) locating a plurality of cells on a support which is capable of maintaining the cells in a consistent environment; b) monitoring a biological agent, or a collection thereof, on or in the cell to produce the profile data for the cell; and

10 c) determining the state of said cell from said data.

53. A method according to Item 52, further comprising the step of correlating said profile and the state of said cell in advance.

15 54. A method according to Item 52, wherein said cell comprises a cell for which the state thereof is known.

20 55. A method according to Item 52, wherein there are at least two types of said biological agent.

56. A method according to Item 52, further comprising the step of arbitrarily selecting said biological agent.

25 57. A method according to Item 52, wherein said data is produced in real-time.

30 58. A method according to Item 52, wherein said status is selected from the group consisting of differentiation state, undifferentiation state, cellular response to a foreign agent, cellular cycle and growth state.

59. A method according to Item 52, wherein said cell is selected from the group consisting of a stem cell and a somatic cell.

5

60. A method according to Item 52, wherein said solid support comprises a substrate.

61. A method according to Item 52, wherein said biological agent is a nucleic acid molecule, and said cell is transfected with said nucleic acid molecule.

10

62. A method according to Item 61, wherein said transfection is conducted on a solid phase or in a liquid phase.

15

63. A method according to Item 52, wherein said biological agent has the capability of binding a different biological agent.

20

64. A method according to Item 52, wherein said step of determination c) comprises comparing the phases of said profile.

65. A method according to Item 52, wherein said step of determination c) comprises obtaining the difference between said profile and a control profile.

25

66. A method according to Item 52, wherein said step of determination c) comprises a mathematical processing method

30

selected from the group consisting of signal processing and multivariate analysis methods.

67. A method for correlating a foreign agent and a cellular  
5 response to the foreign agent, comprising the steps of:

a) subjecting a cell to a foreign agent on a support capable of maintaining a plurality of cell in a consistent environment;

b) monitoring a biological agent or a collection  
10 thereof on or in the cell to produce the profile data for the cell; and

c) correlating the foreign agent and the profile.

68. A method according to Item 68, wherein said cell is  
15 immobilized on said support.

69. A method according to Item 69, further comprising the step of using at least two of said foreign agents to obtain profiles of each of the foreign agents.  
20

70. A method according to Item 67, further comprising the step of classifying at least two of said profiles to classify foreign agents corresponding to the profiles.

25 71. A method according to Item 70, wherein said profile is presented in real-time.

72. A method according to Item 67, wherein said cell is cultured on an array.  
30

73. A method according to Item 67, wherein the monitoring of said profile in step (b) comprises obtaining image data from said array.

5 74. A method according to Item 67, wherein said correlation between said foreign agent and said profile in step (c) is a step of identifying the identity or difference of the phase of said profile.

10 75. A method according to Item 67, wherein said foreign agent is selected from the group consisting of temperature changes, humidity changes, electromagnetic waves, potential difference, visible light, infrared light, ultraviolet light, X-rays, chemical substances, pressure,  
15 gravity changes, gas partial pressure and osmotic pressure.

76. A method according to Item 75, wherein said chemical substance is a biological molecule, a chemically synthesized substance or a culture medium.

20

77. A method according to Item 76, wherein said biological molecule is selected from the group consisting of a nucleic acid, a protein, a lipid, a sugar, a proteolipid, a lipoprotein, a glycoprotein and a proteoglycan.

25

78. A method according to Item 76, wherein said biological molecule comprises at least one biological molecule selected from the group consisting of a hormone, a cytokine, a cell adhesion factor and an extracellular matrix protein.

30

79. A method according to Item 75, wherein said chemical substance is an agonist or antagonist of a receptor.

80. A method for identifying an unidentified foreign agent  
5 given to a cell, from the profile of said cell, comprising the steps of:

a) subjecting a cell on a support capable of maintaining a plurality of cells in a consistent environment, to a plurality of known foreign agents ;

10 b) monitoring a biological agent, or a collection thereof, on or in the cell over time to produce the profile data for the cell in response to each of the known foreign agents to produce profile data for the cell;

15 c) correlating each of the known foreign agents and each of the profiles;

d) subjecting the cell to an unidentified foreign agent;

20 e) monitoring a biological agent, or a collection thereof, on or in the cell subjected to the unknown foreign agent over time to obtain the profile of the cell relating to the unidentified foreign agent;

f) determining the profile corresponding to the profile obtained in step e) amongst the profiles obtained in step b);

25 g) determining that the unidentified foreign agent is the known foreign agent corresponding to the profile which has been determined in step f).

81. A method for identifying an unidentified foreign agent  
30 given to a cell, from the profile of the cell, comprising

the steps of:

a) providing data relating to a correlation between a known foreign agent, and a profile of the cell corresponding to the known foreign agent, with respect to  
5 a biological agent, or a collection thereof, on or in the cell;

b) subjecting the cell to an unidentified foreign agent;

c) monitoring the biological agent, or the collection  
10 thereof, on or in the cell over time to produce a profile of the cell;

d) determining the profile corresponding to the profile obtained in step c) amongst the profiles provided in step a); and

15 e) determining that the unidentified foreign agent is the known foreign agent corresponding to the determined profile.

82. A method for obtaining a profile relating to  
20 information of a cell in a consistent environment, comprising the steps of:

a) locating a plurality of cells to a support which is capable of maintaining the cells in a consistent environment; and

25 b) monitoring a biological agent, or a collection thereof, on or in the cell over time to produce the profile data for the cell.

83. A storage medium on which data produced by a method  
30 according to Item 1, is stored.

84. A storage medium according to Item 83, wherein said storage medium further comprises data of at least one information relating to one selected from the group  
5 consisting of information relating to conditions under said monitoring, information relating to said profile, information relating to the state of said cell and information relating to the biological agent.

10 85. A storage medium according to Item 84, wherein the data is stored in a format which links a plurality of the data to each other.

86. A storage medium according to Item 84, wherein the data  
15 is stored in a format which has links per said cell.

87. Data produced by a method according to Item 1.

88. A transmission medium comprising data produced by a  
20 method according to Item 1.

89. A system for producing profile data relating to information of a cell in a consistent environment, said method comprising:

25 a) a support which is capable of maintaining the cell in a consistent environment;

b) means for monitoring a biological agent, or a collection thereof, on or in the cell to produce the profile data for the cell; and

30 c) means for producing profile data for the cell from



a signal obtained from the means for monitoring.

90. A system according to Item 89, further comprising a plurality of cells, and the plurality of cells are  
5 immobilized on to the support.

91. A system according to Item 90, wherein said support is attached at least one substance selected from the group consisting of a salt and an actin-like substance.  
10

92. A system according to Item 89, wherein said means for monitoring comprises at least one means selected from the group consisting of optical microscopes, fluorescence microscopes, phase-contrast microscopes, reading devices  
15 using a laser source, means using surface plasmon resonance (SPR) imaging, electric signals, chemical or biochemical markers singly or in combination, radiation, confocal microscopes, nonconfocal microscopes, differential interference microscopes, stereoscopic microscopes, video  
20 monitors and infrared cameras.

93. A system for presenting profile data relating to information of a cell in a consistent environment, comprising:

25 a) a support which is capable of maintaining the cell in a consistent environment;

b) means for monitoring a biological agent, or a collection thereof, on or in the cell to produce the profile data for the cell;

30 c) means for producing profile data for the cell from

a signal obtained from the means for monitoring; and  
d) means for presenting the data.

94. A system according to Item 93, further comprising a  
5 plurality of cells, wherein the plurality of cells are  
immobilized on to the support.

95. A system according to Item 93, wherein said support  
is attached at least one substance selected from the group  
10 consisting of a salt and an actin-like substance.

96. A system according to Item 93, wherein said means for  
monitoring comprises at least one means selected from the  
group consisting of optical microscopes, fluorescence  
15 microscopes, phase-contrast microscopes, reading devices  
using a laser source, means using surface plasmon resonance  
(SPR) imaging, electric signals, chemical or biochemical  
markers singly or in combination, radiation, confocal  
microscopes, nonconfocal microscopes, differential  
20 interference microscopes, stereoscopic microscopes, video  
monitors and infrared cameras.

97. A system according to Item 93, wherein said means for  
presenting data is a display.  
25

98. A system according to Item 93, wherein said means for  
presenting data is a speaker.

99. A system for determining the state of a cell in a  
30 consistent environment, comprising

a) a support which is capable of maintaining the cell in a consistent environment;

b) means for monitoring a biological agent, or a collection thereof, on or in the cell to produce the profile data for the cell;

c) means for producing profile data for the cell from a signal obtained from the means for monitoring; and

d) means for determining the state of said cell from said data.

10

100. A system for correlating a foreign agent and a cellular response to the foreign agent, comprising:

a) a support capable of maintaining a plurality of cells in a consistent environment;

b) means for subjecting the cells to a foreign agent;

c) monitoring a biological agent, or a collection thereof, on or in the cells to produce the profile data for the cells;

d) means for producing profile data for the cells from a signal obtained from the means for monitoring; and

e) means for correlating the foreign agent and the profile.

101. A system for identifying an unidentified foreign agent given to a cell, from the profile of said cell, comprising:

a) a support capable of maintaining a plurality of cells in a consistent environment;

b) means for subjecting the cells to a known foreign agent;

c) means for monitoring a biological agent, or a

collection thereof, on or in a cell over time;

d) means for obtaining the profile data for the cell in response to each of known foreign agents to produce profile data for the cell;

5 e) means for correlating each of the known foreign agents and each of the profiles;

f) means for subjecting the cells to an unknown foreign agent;

g) means for comparing the profile of the known foreign agent obtained with means d), and the profile of the unknown foreign agent to determine a profile corresponding to the profile of the unknown foreign agent amongst the profiles of the known foreign agents, wherein said determined unidentified foreign agent is the known foreign agent for  
10 which the determined profile corresponds to.  
15

102. A system for identifying an unidentified foreign agent given to a cell, from the profile of the cell, comprising:

a) a storage medium having stored data relating to a correlation between a known foreign agent, and a profile of the cell corresponding to the known foreign agent, with respect to a biological agent, or a collection thereof, on or in the cell;  
20

b) means for subjecting the cell to an unidentified foreign agent;  
25

c) a support which is capable of maintaining a plurality of cells in a consistent environment;

d) means for monitoring the biological agent or the collection thereof on or in the cell over time to produce a profile of the cell;  
30

e) means for obtaining the profile of the cell from a signal obtained from the means for monitoring; and

f) means for determining the profile corresponding to the profile obtained relating to the unidentified foreign agent amongst the profiles stored in storage medium a),  
5 wherein the unidentified foreign agent is the known foreign agent for which the determined profile corresponds to.

103. A support capable of immobilizing a plurality of cells  
10 and maintaining the cells in a consistent environment.

104. A support according to Item 103, wherein the cells on the support are located in an array format.

15 105. A support according to Item 103 comprising a complex of a positively charged substance and a negatively charged substance; a salt; or an actin-like substance.

20 106. A support according to Item 103 comprising a complex of a positively charged substance and a negatively charged substance; a salt; and an actin-like substance.

25 107. A support according to Item 103, wherein said item is capable of being located within a space of 1 mm or less at maximum.

108. A support according to Item 103, further comprising a cell immobilized thereon.

30 109. A support according to Item 103, further comprising

a biological agent immobilized thereon.

110. A support according to item 109, wherein two or more types of said biological agent are immobilized thereon.

5

111. A support according to Item 103, wherein a cell and a biological agent are immobilized thereon.

112. A support according to Item 103, wherein a salt; a  
10 complex between a positively charged substance and a negatively charged substance; and an actin-like substance are immobilized thereon together with a cell and a biological agent.

15 113. A support according to Item 103, wherein a salt; a complex between a positively charged substance and a negatively charged substance; and an actin-like substance are immobilized thereon together with a cell and a biological agent, in an array format.

20

114. A support according to Item 104, wherein a salt, a gene introduction reagent, and an actin-like substance, a nucleic acid molecule, and a cell are immobilized thereon in an array format.

25

115. A support according to Item 114, wherein the salt is selected from the group consisting of calcium chloride, sodium hydrogen phosphate, sodium hydrogen carbonate, sodium pyruvate, HEPES, calcium chloride, sodium chloride,  
30 potassium chloride, magnesium sulfide, iron nitrate, amino

acids, and vitamins.

116. A support according to Item 114, wherein the gene  
introduction reagent is selected from the group consisting  
5 of cationic polymers, cationic lipids, polyamine-based  
reagents, polyimine-based reagents, calcium phosphate,  
oligofectamine and oligofecter.

117. A support according to Item 114, wherein the  
10 actin-like substance comprises at least one protein selected  
from the group consisting of fibronectin, laminin, and  
vitronectin, or a variant or fragment thereof.

118. A support according to Item 114, wherein the nucleic  
15 acid molecule comprises a sequence encoding a protein  
selected from the group consisting of cytokines, hormones,  
cell adhesion molecules, cytoskeleton proteins and enzymes.

119. A support according to Item 114, wherein the cell  
20 comprises a cell selected from the group consisting of an  
animal cell, an insect cell, a plant cell, a bacterial cell  
and a fungal cell.

120. A support according to Item 114 wherein material of  
25 said support comprises material selected from the group  
consisting of glass, silica and plastics.

121. A method for producing a support comprising a  
plurality of cells immobilized thereon and capable of  
30 maintaining the cells in a consistent environment,

comprising the steps of:

A) providing a support; and

B) immobilizing a cell to the support using a complex comprising of a salt, a positively-charged substance and  
5 a negatively-charged substance.

122. A method according to Item 121, wherein said step of immobilizing comprises immobilizing a mixture of the salt, a gene introduction reagent as the positively-charged  
10 substance, , an actin-like substance, a nucleic acid molecule as the negatively-charged substance, and the cell in an array format.

123. A method according to Item 121, wherein said step of  
15 immobilizing comprises a printing step.

124. A method according to Item 121, wherein the step of providing the support comprises the step of producing the support from a support material.  
20

125. An apparatus for producing a support comprising a plurality of cells immobilized thereon and capable of maintaining the cells in a consistent environment, comprising:

25 A) means for providing a support; and

B) means for immobilizing a cell to the support using a complex comprising a salt, a positively-charged substance and a negatively-charged substance.

30 126. An apparatus according to Item 125, wherein said means



for immobilizing comprises means for printing.

127. An apparatus according to Item 125, wherein the means for providing the support comprises means for shaping the support from a support material.

128. A method for producing a digital cell, comprising the steps of:

a) obtaining a cell parameter specifying a cell of experimental interest;

b) obtaining an environment parameter specifying an environment under which the cell specified by the cell parameter is cultured;

c) obtaining a stimulus parameter specifying a stimulus to be given to the cell specified by the cell parameter;

d) obtaining a stimulus response result showing a result which the cell specified by the cell parameter responds to the stimulus specified by the stimulus parameter under the environment specified by the environment parameter;

e) producing an experimental data for the cell, by correlating the cell parameter, the environment parameter, the stimulus parameter and the stimulus response result; and

f) optionally repeating steps a) through e) to produce at least one collection of experimental data for the cell, and to provide the at least one collection of experimental data as a digital cell.

129. A method according to Item 128, wherein the environment parameter comprises a parameter indicating culture medium in which the cell is cultured, and a parameter showing the conditions of the culture medium.

5

130. A method according to Item 128, wherein the stimulus parameter comprises a parameter showing a reporter and a parameter showing a chemical stimulus.

10 131. A method according to Item 128, wherein said stimulus response result comprises profile data for the cell obtained by monitoring a biological agent, or a collection thereof, on or in the cell over time.

15 132. A method according to Item 128, further comprising the step of storing the digital cell to a database.

133. An apparatus for producing a digital cell, comprising:

20 a) means for obtaining a cell parameter specifying a cell of experimental interest;

b) means for obtaining an environment parameter specifying an environment under which the cell specified by the cell parameter is cultured;

25 c) means for obtaining a stimulus parameter specifying a stimulus to be given to the cell specified by the cell parameter;

30 d) means for obtaining a stimulus response result showing a result which the cell specified by the cell parameter responds to the stimulus specified by the stimulus parameter under the environment specified by the environment

parameter;

e) means for producing an experimental data for the cell, by correlating the cell parameter, the environment parameter, the stimulus parameter and the stimulus response  
5 result; and

f) means for providing at least one collection of experimental data as a digital cell, by optionally repeating steps performed by the means a) through e) to produce at least one collection of experimental data for the cell.  
10

134. A method for providing a service which reproduces an experimental result of an actual cell using a digital cell by means of a computer system comprising a service requester and a service provider, comprising the steps of:

15 preparing a database having at least one digital cell stored thereon, wherein the at least one digital cell is expressed as a collection of at least one experimental data of a cell of experimental interest, wherein each of the at least one experimental data comprises a cell parameter  
20 specifying the cell, an environment parameter specifying an environment under which the cell specified by the cell parameter is cultured, a stimulus parameter specifying a stimulus to be given to the cell specified by the cell parameter, and a stimulus response result showing a result  
25 which the cell specified by the cell parameter responds to the stimulus specified by the stimulus parameter under the environment specified by the environment parameter;

receiving the cell parameter, the environment parameter and the stimulus parameter by the service  
30 requester to produce a request comprising the cell parameter,

the environment parameter and the stimulus parameter;

providing the request to the service provider by the service requester;

searching the database in response to the request by  
5 the service provider to determine whether or not there is  
the stimulus response result relating to the cell parameter,  
the environment parameter and the stimulus parameter  
included in the request, in the database;

providing the stimulus response result to the service  
10 requester by the service provider, when it is determined  
that the stimulus response result relating to the cell  
parameter, the environment parameter and the stimulus  
parameter included in the request exists in the database;  
and

15 presenting the stimulus response result by the service  
requester.

135. A method for providing a service for reproducing an  
experimental result of an actual cell using a digital cell,  
20 by means of a computer system comprising a service requester  
and a plurality of service providers, comprising the steps  
of:

preparing a plurality of databases, each having at  
least one digital cell stored thereon, wherein the at least  
25 one digital cell is expressed as a collection of at least  
one experimental data of a cell of experimental interest,  
wherein each of the at least one experimental data comprises  
a cell parameter specifying the cell, an environment  
parameter specifying an environment under which the cell  
30 specified by the cell parameter is cultured, a stimulus

parameter specifying a stimulus to be given to the cell specified by the cell parameter, and a stimulus response result showing a result which the cell specified by the cell parameter responds to the stimulus specified by the stimulus parameter under the environment specified by the environment parameter;

preparing a service registry which stores at least one service capable of being provided by the plurality of service providers;

10 receiving the cell parameter, the environment parameter and the stimulus parameter by the service requester to produce a request comprising the cell parameter, the environment parameter and the stimulus parameter;

searching the service registry in response to the request by the service requester to determine whether or not there exists a service provider capable of providing a service for the request amongst the plurality of service providers;

providing the request to the service provider by the service requester when it is determined that a service provider capable of providing a service of the request amongst the plurality of service providers exists;

searching the database in response to the request by the service provider to determine whether or not there is the stimulus response result relating to the cell parameter, the environment parameter and the stimulus parameter included in the request in the database;

providing the stimulus response result to the service requester by the service provider, when it is determined that the stimulus response result relating to the cell

parameter, the environment parameter and the stimulus parameter included in the request exists in the database; and

presenting the stimulus response result by the service requester.

136. A computer system for providing a service which reproduces an experimental result of an actual cell using a digital cell, comprising:

10 a service requester being composed such that it can have access to a database having at least one digital cell stored thereon, each of the at least one digital cell is expressed as a collection of at least one experimental data of a cell of experimental interest, wherein each of the at least one experimental data comprises a cell parameter specifying the cell, an environment parameter specifying an environment under which the cell specified by the cell parameter is cultured, a stimulus parameter specifying a stimulus to be given to the cell specified by the cell parameter, and a stimulus response result showing a result which the cell specified by the cell parameter responds to the stimulus specified by the stimulus parameter under the environment specified by the environment parameter; and

15 a service provider requesting a service desired by a user;

25 wherein the service requester comprises:

means for receiving the cell parameter, the environment parameter and the stimulus parameter to produce a request comprising the cell parameter, the environment parameter and the stimulus parameter; and

means for providing the request to the service provider, and

wherein the service provider comprises:

means for searching the database in response to  
5 the request by the service provider to determine whether  
or not there is the stimulus response result relating to  
the cell parameter, the environment parameter and the  
stimulus parameter included in the request in the database;  
and

10 means for providing the stimulus response  
result to the service requester by the service provider,  
when it determined that the stimulus response result  
relating to the cell parameter, the environment parameter  
and the stimulus parameter included in the request exists  
15 in the database;

wherein the service requester further comprises

means for presenting the stimulus response  
result by the service requester.

20 137. A computer system according to Item 136 wherein the  
service requester is a Web browser which the user operates,  
and the service provider is a Web server linked to the service  
requester via the Internet.

25 138. A computer system according to Item 136, wherein the  
service requester provides the request to the service  
provider in a format described in XML language.

139. A computer system according to Item 136, wherein the  
30 service provider provides the stimulus response result to

the service requester in a format described in XML language.

140. A computer system for providing a service which reproduces an experimental result of an actual cell using  
5 a digital cell, comprising:

a plurality of service providers, each composed such that the plurality of service providers can have access to a database having at least one digital cell stored thereon, each of the at least one digital cell is expressed as a  
10 collection of at least one experimental data of a cell of experimental interest, wherein each of the at least one experimental data comprises a cell parameter specifying the cell, an environment parameter specifying an environment under which the cell specified by the cell parameter is  
15 cultured, a stimulus parameter specifying a stimulus to be given to the cell specified by the cell parameter, and a stimulus response result showing a result which the cell specified by the cell parameter responds to the stimulus specified by the stimulus parameter under the environment  
20 specified by the environment parameter;

a service registry which stores at least one service which the plurality of service providers can provide; and

a service provider requesting a service desired by a user;

25 wherein the service requester comprises:

means for receiving the cell parameter, the environment parameter and the stimulus parameter to produce a request comprising the cell parameter, the environment parameter and the stimulus parameter;

30 means for searching the service registry in



response to the request by the service requester to determine whether or not there exists a service provider capable of providing a service of the request amongst the plurality of service providers and

5                   means for providing the request to the service provider by the service requester when it is determined that there exists a service provider capable of providing a service of the request amongst the plurality of service providers,

10           wherein each of the plurality of service providers comprises:

                  means for searching the database in response to the request by the service provider to determine whether or not there is the stimulus response result relating to  
15 the cell parameter, the environment parameter and the stimulus parameter included in the request in the database;  
and

                  means for providing the stimulus response result to the service requester by the service provider,  
20 when it is determined that the stimulus response result relating to the cell parameter, the environment parameter and the stimulus parameter included in the request exists in the database;

                  wherein the service requester further comprises  
25                   means for presenting the stimulus response result by the service requester.

141. A computer system according to Item 140, wherein the service requester is a Web server connected to a Web browser  
30 which the user operates via the Internet, and each of the

plurality of service providers is a Web server connected to the service requester via the Internet.

142. A computer system according to Item 140, wherein the  
5 service requester provides the request to the service provider in a format described in XML language.

143. A computer system according to Item 140, wherein the  
service provider provides the stimulus response result to  
10 the service requester in a format described in XML language.

144. A method for producing the profile data relating information of a cell, comprising the steps of:

a) immobilizing and locating a cell on a support; and  
15 b) monitoring a biological agent, or a collection thereof, on or in the cell to produce the profile data for the cell.

145. A method according to Item 144, wherein the biological  
20 agent is a nucleic acid molecule or a molecule derived from the nucleic acid molecule.

146. A method according to Item 144, wherein the cell is immobilized to the support by a composition comprising a)  
25 a complex with a positively charged substance and a negatively charged substance; and b) a salt.

147. A method according to Item 144 wherein the cell is provided with an actin-like substance.

30

148. A method according to Item 144, wherein the cell is immobilized to the support by a composition comprising a) a complex with a positively charged substance and a negatively charged substance; and b) a salt, and is provided  
5 with an actin-like substance.

149. A method according to Item 144 wherein the biological agent is selected from the group consisting of a nucleic acid molecule, a protein, a saccharide, a lipid, a low  
10 molecule, and a complex thereof.

150. A method according to Item 144, wherein the cell is cultured for at least about three days before the step of monitoring.  
15

151. A method according to Item 144, wherein the biological agent comprises a nucleic acid molecule encoding a gene.

152. A method according to Item 144, wherein the profile  
20 comprises a profile of gene expression.

153. A method according to Item 144, wherein the profile comprises a profile of an apoptosis signal.

25 154. A method according to Item 144 wherein the profile is a profile of a stress signal.

155. A method according to Item 144 wherein the profile is a profile of the localization of a molecule.  
30

156. A method according to Item 155 wherein the molecule is detected by means selected from the group consisting of fluorescence, phosphorescence, radioactivity, and a combination thereof.

5

157. A method according to Item 144 wherein the profile comprises a variation of cell morphology.

158. A method according to Item 144 wherein the profile  
10 comprises a profile of promoters.

159. A method according to Item 144, wherein said profile comprises a profile of a promoter dependent on a specific drug.

15

160. A method according to Item 144 wherein said profile comprises a profile of a promoter dependent on a specific drug, wherein said method further comprises the step of administering the specific drug.

20

161. A method according to Item 144 further comprising the step of subjecting the cell to a foreign agent.

162. A method according to Item 161, wherein said foreign  
25 agent comprises an RNAi.

163. A method according to Item 161, wherein said foreign agent comprises a chemical not present in a biological body.

30 164. A method according to Item 144, wherein said profile

comprises a profile of intermolecular interaction.

165. A method according to Item 161, wherein said foreign agent comprises a ligand for a receptor of said cell.

5

166. A method according to Item 144, wherein said profile comprises a profile of an interaction between a receptor and a ligand.

10 167. A method according to Item 144, wherein said profile is a cellular form, and said method further comprises the step of giving to said cell a stimulus selected from the group consisting of overexpression, underexpression or knockdown of a gene, addition of a foreign agent and a change  
15 in the environment.

168. A method according to Item 144, wherein said profile comprises a profile of interaction between molecules present in said cell.

20

169. A method according to Item 144, further comprising the step of conducting observation using a technology selected from the group consisting of two-hybrid method, FRET and BRET.

25

170. A method according to Item 144, wherein said profile comprises a profile of interaction between molecules present in said cell, wherein the method further comprises the step of conducting observation using a technology selected from  
30 the group consisting of two-hybrid method, FRET and BRET.

171. A method according to Item 144, wherein said cell is located on said support in an array format.

5 172. A method according to Item 144, wherein said cell is located on said support in an array format, and each of said plurality of cells are located at a space of 1 mm at maximum.

173. A method according to Item 144, wherein said profile  
10 is obtained at real time.

174. A method according to Item 144 further comprising the step of immobilizing said cell to a solid support.

15 175. A method according to Item 144, wherein said data comprises information relating to said profile.

176. A method according to Item 144 wherein said data comprises information relating to conditions during said  
20 monitoring.

177. A method according Item 144 wherein said data comprises information relating to the state of said cell.

25 178. A method according to Item 144 wherein said biological agent to be monitored comprises at least two types of biological agent.

179. A method according to Item 144 wherein said biological  
30 agent to be monitored comprises at least three types of

biological agent.

180. A method according to Item 144 wherein said biological agent to be monitored comprises at least eight types of  
5 biological agent.

181. A method according to Item 144 further comprising the step of arbitrarily selecting a biological agent.

10 182. A method according to Item 144 wherein said cell is selected from the group consisting of a stem cell and a somatic cell.

183. A method according to Item 144 wherein said support  
15 comprises a solid support.

184. A method according to Item 144 wherein said support comprises a substrate.

20 185. A method according to Item 144 wherein said biological agent is a nucleic acid molecule, and said cell is transfected with said nucleic acid molecule.

186. A method according to Item 185, wherein said  
25 transfection is conducted on a solid phase or in a liquid phase.

187. A method according to Item 185, wherein said transfection is conducted on a solid support.

30

188. A method according to Item 144, further comprising the step of comparing a phase of said profile.

189. A method according to Item 144, further comprising the  
5 step of subtracting a control profile from the profile of said cell.

190. A method according to Item 144 further comprising the step of processing the profile with a mathematical  
10 processing method selected from signal processing and multivariant analysis methods.

Hereinafter, the present invention will be described by way of preferred embodiments. It will be  
15 understood by those skilled in the art that the embodiments of the present invention can be appropriately made or carried out based on the description of the present specification and the accompanying drawings, and commonly used techniques well known in the art. The function and effect of the present  
20 invention can be easily recognized by those skilled in the art.

The present invention provides consecutive information (profile) data relating to the state of a cell.  
25 The present invention provides information and data (in particular consecutive information and consecutive profile) relating to the state of a cell in a consistent environment, in a reproducible manner. According to the present invention, a method and a system for presenting such  
30 data in an accurate manner are provided. Specifically, it



should be surprising a effect to be able to provide a system and a method for information at the cellular level in a consistent environment in terms of a complex system as such or in a direct manner, and to be able to provide such data and technology for aligning such data *per se*. The present invention further attains the effect of providing a digital cell based on actual, live data and the use thereof, which have not been conventionally possible prior to the present invention.

As such, according to the present invention, it is possible to determine the state of cells by observing a surprisingly small number of factors. Therefore, the present invention is applicable to diagnosis, prevention, and treatment. The present invention is also applicable to the fields of food, cosmetics, agriculture, environmental engineering, and the like. Reproduction of a live experiment on a computer system attains the effect of enabling education and research in the field of biotechnology.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the results of experiments in which various actin-like substances and HEK293 cells were used, where gelatin was used as a control. Figure 1 shows the effect of each adhesion substance (HEK cell) with respect to transfection efficiency. The HEK cells were transfected with pEGFP-N1 using an Effectene reagent.

Figure 2 shows exemplary transfection efficiencies when fibronectin fragments were used.

5                Figure 3 shows exemplary transfection efficiencies when fibronectin fragments were used.

                 Figure 4 shows a summary of the results presented in Figures 2 and 3.

10

                 Figure 5 shows the results of an example in which transfection efficiency was studied for various cells.

                 Figure 6 shows the results of transfection when  
15 various plates were used.

                 Figure 7 shows the results of transfection when various plates were used at a fibronectin concentration of 0, 0.27, 0.53, 0.8, 1.07, and 1.33 ( $\mu\text{g}/\mu\text{L}$ ).

20

                 Figure 8 provides exemplary photographs showing cell adhesion profiles in the presence or absence of fibronectin.

25                Figure 9 shows exemplary cross-sectional photographs of cell adhesion profiles in the presence or absence of fibronectin. Cross-sections of human mesenchymal stem cells (hMSC) were observed using a confocal laser scanning microscope. hMSC were stained with SYTO61  
30 (blue fluorescence) and Texas red - X phalloidin (red

fluorescence) and fixed with 4% PFA. Blue fluorescence (nuclei: SYTO61) and red fluorescence (nuclei: Texas red - X phalloidin) were obtained using a confocal laser microscope (LSM510, Carl Zeiss Co., Ltd., pin hole size=1.0, image interval=0.4).

Figure 10 shows transition of nuclear surface area. Relative nuclear surface area was determined by cross-sections of hMSC observed with confocal laser scanning microscopy. hMSC was fixed with 4% PFA.

Figure 11 shows the results of an exemplary transfection experiment when a transfection array chip was constructed and used.

Figure 12 shows exemplary contamination between each spot on an array.

Figure 13 shows an experiment in which spatially-spaced DNA was taken into cells after the solid phase transfection of the present invention in Example 4.

Figure 13A schematically shows a method for producing a solid phase transfection array (SPTA). This figure shows the methodology of a solid transfection.

Figure 13B shows the results of a solid phase transfection. A HEK293 cell line was used to produce a SPTA. Green colored portions indicate transfected adherent cells. According to this result, the method of the present invention

can be used to produce a group of cells separated spatially and transfected with different genes. As such, Figure **13A-B**, as a whole, depicts schematically the procedure of transfection (SPTA). Figure **13A** depicts the outlines of SPTA determination, and Figure **13B** depicts a result of SPTA using HEK293 cell strain. The bar indicates 3mm.

Figure **13C** shows the difference between conventional liquid phase transfection and SPTA.

Figure **14** shows the results of comparison of liquid phase transfection and SPTA.

Figure **14A** shows the results of experiments where 5 cell lines were measured with respect to GFP intensity/mm<sup>2</sup>. Transfection efficiency was determined as fluorescence intensity per unit area.

Figure **14B** shows fluorescent images of cells expressing EGFP corresponding to the data presented in Figure **14A**. White circular regions therein were regions in which plasmid DNA was fixed. In other regions, cells were also fixed in solid phase, however, cells expressing EGFP were not observed. The white bar indicates 500  $\mu$ m.

Fluorescent photographs of EGFP expressing cells corresponding to Figure **14A** are shown with respect to the five types of cells per measured fluorescence/mm<sup>2</sup>. White circles correspond to plasmid DNA printed regions. Outside

these regions, cells express EGFP. Further, regions other than the printed regions are attached cells.

Figure **14C** shows an exemplary transfection  
5 method of the present invention.

Figure **14D** shows an exemplary transfection method of the present invention.

10 Figure **15** shows the results of coating a chip, whereby cross contamination was reduced.

Figure **15** shows the results of liquid phase transfection and SPTA using HEK293 cells, HeLa cells, NIT3T3  
15 cells (also referred to as "3T3"), HepG2 cells, and hMSCs. Transfection efficiency was determined by GFP intensity.

Transfection efficiency of hMSC depending on the N/P ratio used is shown in Figure **15A**. In the phase of prior  
20 liquid phase transfection (Figure **15B**, upper panel), hMSC cells were dead and in the case of SPTA, cell morphology was normal (Figure **15B**, lower panel).

Figure **16** shows cross contamination between  
25 each spot. A nucleic acid mixture containing fibronectin having a predetermined concentration was fixed to a chip coated with APS or PLL (poly-L-lysine). Cell transfection was performed on the chip. Substantially no cross contamination was observed (upper and middle rows). In  
30 contrast, significant cross contamination of fixed nucleic

acids was observed on an uncoated chip (lower row).

Figure **16C** shows a correlation relationship between the types of substances contained in a mixture used for fixation of nucleic acid and the cell adhesion rate. The graph presented in Figure **16** shows an increase in the proportion of adherent cells over time. A longer time is required for cell adhesion when the slope of the graph is shallow than when the slope of the graph is steep.

Figure **16D** is an enlarged graph which is presented in Figure **16C**.

Figure **17** shows an exemplary configuration of a computer which was used to perform the method of the present invention.

Figure **18** depicts an example of a mathematical analysis method of the present invention. Profiles of promoters shown in Figure **18A** (average of pNEFAT-d2EGFP/negative control) and Figure **18B** (average of pERE-d2EGFP/negative control) are obtained by measuring the fluorescent intensity thereof over time. These profiles have been normalized using the autologous fluorescence of either the cell or medium used. Thereafter, in order to compare amplitude of the reporter expression fluctuation, an amplitude = 5 or more ( $TH \geq 5$ ) was determined to have change in expression fluctuation state. Further, differentiation induction was divided into the following sections: start of differentiation induction, early stage (0-17.5 hours),

and late stage (17.5-31.5 hours) and total stages (0-31.5 hours); and those observed with a variation in expression of an amplitude of 5 or more ( $TH \geq 5$ ) were defined as (+) and those with an amplitude of less than 5 were defined as (-). Based on these definitions, the profiles of A and B were evaluated as shown in the lower tables of Figures **18A** and **18B**. In the table, when extracting any number of reporters, (A+B+ ... n) have been integrated with respect to n types of wave forms and the sum is divided by n to form the average wave form and if variations beyond the threshold, such variations were deemed as being "changed".

Figure **18B** depicts another example of a mathematical analysis according to the present invention. When a reporter is extracted (A+B+ ... n), n types of wave types are integrated, and the sum is divided by n to produce an average wave form, which was deemed as being a change of the variation above a threshold. The left hand panel of Figure **18B** depicts the integration of two reporter profiles and draws the average wave form in red or with solid squares. Those with 5 or more variations of the average profile were deemed to be expression variations for evaluation. As a result, evaluation can be conducted for variation of the two extracted reporters, as shown in the table herein.

25

Figure **19** depicts exemplary plasmids containing promoters used in the present invention and an analysis according to the present invention. Seventeen types of transcriptional factors shown in the left hand panel of Figure 19 were used as a reporter under the conditions of

30

osteoblast differentiation and maintenance of an undifferentiated mesenchymal stem cell, and the expression profile thereof have been obtained over time (Figure 19, right handed panel). From these seventeen types of profiles, any number of profiles have been extracted and evaluated by the method as previously described in Figure 18, taking the change in amplitude of the response profile of each transcriptional factor as a standard.

Figure 20 depicts an example of mathematical analysis at the early stage of induction of differentiation. By changing the combination arbitrarily extracted in the early differentiation induction stage, results as shown in Figure 20 have been obtained. Any number of reporters were extracted from the reporter group consisting of seventeen species, and calculated for the average profile according to the method shown in Figure 18. Those having five or more variation widths are the results evaluated with the evaluation windows 0-31.5, 0-17.5 and 17.5-31.5. Each extraction condition has seventeen extraction patterns, except for where the seventeen extraction pattern has only one way of extraction. Amongst these combinations, Figure 20 shows the ratio in which variation is found therein, including the table and graph included therein. This analysis allows confirmation of differentiation after fifteen hours although it is not possible to understand the very early stages of differentiation. The number of extraction where 100 % change is found for variation is eight or more in this instance.



Figure **21** depicts an example of a mathematical analytical result at the undifferentiation maintenance stage. As in Figure 20, similar results as shown in the graphs have been obtained when a combination arbitrarily  
5 extracted under conditions to maintain undifferentiation. Comparing the results with the stage of differentiation induction, as in Figure 20, the results are dramatically different. Based on this comparison, it is believed that it is possible to determine whether a cell is moving in the  
10 direction of cell differentiation induction, or instead maintaining an undifferentiated state.

Figure **22** schematically shows a cocktail party process.  
15

Figure **23** shows an exemplary construct of a gene transcription switch reporter used in a transfection plasmid of the present invention.

Figure **24** shows exemplary construction of a set of transcription factor reporters.  
20

Figure **25** shows the results of exemplary assays using transcription factor reporters.  
25

Figure **26** shows an example of measurements of transcriptional activity in the bone differentiation process, , taken in a time-series manner.

Figure **27** shows an example of the oscillation  
30

phenomenon and phase analyses of transcriptional activity.

Figure **28** shows a protocol of siRNA experiment.

5                    Figure **29A** shows the results of the siRNA experiments. The upper panel shows the results of hMSC, and the lower panel shows the results of HeLa cells. The numerals show the concentrations ( $\mu\text{g}/\mu\text{L}$ ) of the siRNA used. The results obtained with the anti-GFP siRNA are shown on  
10 the left hand side, and the right hand side shows the results with the scramble siRNAs.

                  Figure **29B** shows the effects of siRNA when solid transfection (PC12) was conducted on a collagen IV coating.  
15 Figure **29B(A)** shows PC12 cells cotransfected with EGFP vector and anti-EGFP siRNA. As shown, it was observed that only HcRed was colored, and green signals derived from pEGFP-N1 were suppressed. On the other hand, Figure **29B(B)** shows an example using scramble siRNA. As shown, green  
20 fluorescence was observed and thus the effects observed in Figure **29B(A)** are due to the effects of RNAi. Figures showing the relative fluorescence intensities in Figures **29B(A)** and **29B(B)**, are summarized in Figure **29B(C)**. The y axis indicates relative intensity. It can be seen that  
25 effects induced by EGFP were almost completely suppressed.

                  Figure **29C** depicts results and graphs summarizing the above. The left-hand panel is a photograph comparing EGFP RNAi and scramble (mock) RNAi when changing  
30 the ratio of RNAi and pRNA. As shown, EGFP RNAi showed

inhibitory effects, whereas scramble RNAi did not exert such effects. This is shown in the right-hand panel, together with DsRed2. Experimental conditions were in accordance with those described herein. As a result, red (DsRed  
 5 derived signal) and green (EGFP derived signal) were found to be in proportion with the effect of RNAi.

Figure **29D** depicts an exemplary chip used in the RNAi reporter. When using RNAi as input signals and  
 10 coinroducing a gene product capable of transmitting signals, such as EGF and the like, together with a nucleic acid encoding a gene of interest (including a promoter), observation of such signal transmission as output allows extraction of cell information.

15 Figure **29E** shows an exemplary experiment using a variety of reporters (pAP1-EGFP, pAP1(PMA)-EGFP, pCRE-EGFP, pE2F-EGFP, pERE-EGFP, pGAS-EGFP, pGRE-EGFP, pHSE-EGFP, pISRE-EGFP, pMyc-EGFP, pNFAT-EGFP, pNFkB-EGFP,  
 20 pRARE-EGFP, pRb-EGFP, pSTST3-EGFP, pSRE-EGFP, pTRE-EGFP, pp53-EGFP, pCREB-sensor, pIkB-sensor, pp53-sensor, pCasapase3-sensor); the is-element sequence was purchased from Clontech using a plasmid vector produced by recombining a fluorescence protein gene).

25 Figure **30** shows changes in the profile when using tetracycline dependent promoters.

Figure **31** shows expression when using  
 30 tetracycline dependent promoters and tetracycline

independent promoters.

Figure **31B** shows an exemplary result of analysis using a transfected microarray with respect to the effects of tyrosine kinase RNAi on neurons.

Figure **31C** depicts responses to retinoic acid (RA) and nerve growth factor (NGF) by a variety of tyrosine kinases. Inhibition percent by siRNA is shown.

10

Figure **31D** depicts an example of a signaling pathway obtained as a result of an analysis.

Figure **31E** shows the results obtained by the above-mentioned analysis. It shows a general analysis of the tyrosine kinases responsible for human neuron differentiation. Classification is conducted by determining whether it is dopaminergic neuron, cholinergic neuron, or both, or neither. It can be concluded by the analysis that there is high possibility that those tyrosine kinases relating to both types of neuron are involved in neuron projection formation.

Figure **31F** depicts an example of real-time monitoring of transcription regulation of apoptosis in a HeLa cell. The left handed panel shows the result over time, and the right handed panel shows the result of a signaling pathway based on the analysis thereof.

Figure **32** depicts an example of a system

configuration.

Figure **33A** depicts an example of a digital cell according to the present invention.

5

Figure **33B** depicts another example of a digital cell according to the present invention.

Figure **34** depicts an example of a method for  
10 producing a digital cell according to the present invention.

Figure **35** depicts an example of a configuration of computer system 3501 which provides a service reproducing an experimental result obtained using an actual cell using  
15 the digital cell.

Figure **36** depicts an example of procedures of a process which provides a service reproducing an experimental result obtained using an actual cell using the  
20 digital cell.

Figure **37** depicts an example of input interface for inputting cell parameters, environment parameters and stimulus parameters into service requester 3510.

25

Figure **38** depicts an example of configurations of computer system 3801 for providing a service of reproducing an experimental result against an actual cell using the digital cell.

30

Figure 39 depicts an example of procedures of a process for providing a service of reproducing an experimental result against an actual cell using the digital cell.

5

#### DESCRIPTION OF SEQUENCE LISTING

SEQ ID NO.: 1: a nucleic acid sequence encoding  
 10 fibronectin (human)  
 SEQ ID NO.: 2: an amino acid sequence of  
 fibronectin (human)  
 SEQ ID NO.: 3: a nucleic acid sequence encoding  
 vitronectin (mouse)  
 15 SEQ ID NO.: 4: an amino acid sequence of  
 vitronectin (mouse)  
 SEQ ID NO.: 5: a nucleic acid sequence encoding  
 laminin (mouse  $\alpha$ -chain)  
 SEQ ID NO.: 6: an amino acid sequence of laminin  
 20 (mouse  $\alpha$ -chain)  
 SEQ ID NO.: 7: a nucleic acid sequence encoding  
 laminin (mouse  $\beta$ -chain)  
 SEQ ID NO.: 8: an amino acid sequence of laminin  
 (mouse  $\beta$ -chain)  
 25 SEQ ID NO.: 9: a nucleic acid sequence encoding  
 laminin (mouse  $\gamma$ -chain)  
 SEQ ID NO.: 10: an amino acid sequence of  
 laminin (mouse  $\gamma$ -chain)  
 SEQ ID NO.: 11: an amino acid sequence of  
 30 fibronectin (bovine)

SEQ ID NO.: 12: siRNA used in the Examples

SEQ ID NO.: 13: mouse olfactory receptor I7  
(heptanal-sensitive) nucleic acid (Genbank Accession  
No. AF106007)

5 SEQ ID NO.: 14: amino acid sequence of the  
protein encoded by the nucleic acid set forth in SEQ ID  
NO.: 13

SEQ ID NO: 15: nucleic acid encoding the murine  
olfactory receptor S1 (mc9/bc9-equi-sensitive) (Genbank  
10 Accession Number AF121972)

SEQ ID NO: 16: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 15

SEQ ID NO: 17: nucleic acid encoding the murine  
olfactory receptor S50 (cc9-sensitive) (Genbank Accession  
15 Number AF121980)

SEQ ID NO: 18: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 17

SEQ ID NO: 19: nucleic acid encoding the murine  
olfactory receptor S19 (mc9/mh9/bc9-equi-sensitive)  
20 (Genbank Accession Number AF121976)

SEQ ID NO: 20: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 19

SEQ ID NO: 21: nucleic acid encoding the murine  
OR23 (lyral-sensitive) (only coding region of Genbank  
25 Accession Number X92969)

SEQ ID NO: 22: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 21

SEQ ID NO: 23: nucleic acid encoding the murine  
olfactory receptor mOR-EV (vanillin-sensitive) (Genbank  
30 Accession Number AB061229)

SEQ ID NO: 24: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 23

SEQ ID NO: 25: nucleic acid encoding the murine  
olfactory receptor or37a (Genbank Accession Number  
5 AJ133424)

SEQ ID NO: 26: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 25

SEQ ID NO: 27: nucleic acid encoding the murine  
olfactory receptor C6 (Genbank Accession Number AF102523)

10 SEQ ID NO: 28: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 27

SEQ ID NO: 29: nucleic acid encoding the murine  
olfactory receptor F5 (Genbank Accession Number AF102531)

15 SEQ ID NO: 30: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 29

SEQ ID NO: 31: nucleic acid encoding the murine  
olfactory receptor S6 (Genbank Accession Number AF121974)

SEQ ID NO: 32: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 31

20 SEQ ID NO: 33: nucleic acid encoding the murine  
olfactory receptor S18 (Genbank Accession Number AF121975)

SEQ ID NO: 34: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 33

25 SEQ ID NO: 35: nucleic acid encoding the murine  
olfactory receptor S25 (Genbank Accession Number AF121977)

SEQ ID NO: 36: amino acid sequence of the protein  
encoded by the nucleic acid set forth in SEQ ID NO: 35

SEQ ID NO: 37: nucleic acid encoding the murine  
olfactory receptor S46 (Genbank Accession Number AF121979)

30 SEQ ID NO: 38: amino acid sequence of the protein



encoded by the nucleic acid set forth in SEQ ID NO: 37

SEQ ID NO: 39: nucleic acid encoding the  $\alpha$  subunit of murine G-coupled protein (Genbank Accession Number M36778)

5 SEQ ID NO: 40: amino acid sequence of the protein encoded by the nucleic acid set forth in SEQ ID NO: 39

SEQ ID NO: 41: nucleic acid encoding the  $\beta$  subunit of murine G-coupled protein (Genbank Accession Number M87286)

10 SEQ ID NO: 42: amino acid sequence of the protein encoded by the nucleic acid set forth in SEQ ID NO: 41

SEQ ID NO: 43: nucleic acid encoding the  $\gamma$  subunit of murine G-coupled protein (Genbank Accession Number U37527)

15 SEQ ID NO: 44: amino acid sequence of the protein encoded by the nucleic acid set forth in SEQ ID NO: 43

SEQ ID NO: 45: nucleic acid encoding the epidermal growth factor receptor (Genbank Accession Number BC023729)

20 SEQ ID NO: 46: amino acid sequence of the protein encoded by the nucleic acid set forth in SEQ ID NO: 45

SEQ ID NO: 47: the sequence of siRNA used in Example 9

25 SEQ ID NO: 48: the sequence of scrambled RNA used in Example 9

#### BEST MODE FOR CARRYING OUT THE INVENTION

Hereinafter, the present invention will be described. It should be understood throughout the present specification that articles for a singular form (e.g., "a",

30

"an", "the", etc. in English) include the concept of their plurality unless otherwise mentioned. It should be also understood that the terms as used herein have definitions typically used in the art unless otherwise mentioned.

5 Accordingly, unless otherwise defined, all technical and scientific terms used herein shall have the same meaning as generally understood by those skilled in the art to which the present invention pertains. If there is any inconsistency, the present specification precedes,

10 including definitions.

(Definition of terms)

Terms particularly used herein are defined as follows.

15

(Cellular Biology)

The term "cell" is herein used in its broadest sense in the art, referring to a structural unit of the tissue of a multicellular organism, which is capable of self replicating, has genetic information and a mechanism for

20 expressing it, and is surrounded by a membrane structure which isolates the cell from the outside. Cells used herein may be either naturally-occurring cells or artificially modified cells (e.g., fusion cells, genetically modified

25 cells, etc.), as long as the cell has a chemical receptor or is capable of having such a chemical receptor introduced therein. Examples of cell sources include, but are not limited to, a single-cell culture; the embryo, blood, or body tissue of normally-grown transgenic animals; a mixture

30 of cells derived from normally-grown cell lines; and the

like.

As used herein, the term "digital cell" refers to a collection of at least one experimental data on a cell of experimental interest. These experimental data correlate the experimental conditions and the experimental results of an example conducted against an actual cell. The digital cell is constituted such that once an experimental condition is given, the experimental result related to said experimental condition will be reproduced. The digital cell contemplated by the present invention comprises any cell which is amenable to an experiment. It should be understood that the description with respect to all the (living) cells described herein can be applied to a digital cell according to the present invention, as long as such description is applicable to the digital cell.

Using digital cells of the present invention allows reproduction of an experimental result of an experiment conducted using an actual cell, in a computer system. As such, the present invention allow research institutes, educational organizations and individuals having no experimental facilities, to conduct education and advanced research relating to a cell. As a result, business entities in different fields will be able to start business in this field, which has not been possible to date.

Cells used herein may be derived from any organism (e.g., any unicellular organism (e.g., bacteria and yeast) or any multicellular organisms (e.g., animals

(e.g., vertebrates and invertebrates), plants (e.g., monocotyledons and dicotyledons, etc.)). For example, cells used herein are derived from a vertebrate (e.g., Myxiniiformes, Petronyzoniformes, Chondrichthyes, 5 Osteichthyes, amphibian, reptilian, avian, mammalian, etc.), more preferably mammalian (e.g., monotremata, marsupialia, edentate, dermoptera, chiroptera, carnivore, insectivore, proboscidea, perissodactyla, artiodactyla, tubulidentata, pholidota, sirenia, cetacean, primates, 10 rodentia, lagomorpha, etc.). In one embodiment, cells derived from primates (e.g., chimpanzee, Japanese monkey, human) are used. Particularly, without limitation, cells derived from a human are used. The above-described cells may be either stem cells or somatic cells. Also, the cells 15 may be adherent cells, suspended cells, tissue forming cells, and mixtures thereof. The cells may be used for transplantation.

Any organ may be targeted by the present 20 invention. A tissue or cell targeted by the present invention may be derived from any organ. As used herein, the term "organ" refers to a morphologically independent structure, localized to a particular portion of an individual organism, in which a certain function is 25 performed. In multicellular organisms (e.g., animals, plants), an organ consists of several tissues spatially arranged in a particular manner, each tissue being composed of a number of cells. An example of such an organ includes an organ relating to the vascular system. In one embodiment, 30 organs targeted by the present invention include, but are

not limited to, skin, blood vessels, cornea, kidney, heart, liver, umbilical cord, intestine, nerve, lung, placenta, pancreas, brain, peripheral limbs, retina, and the like. As used herein, cells differentiated from a pluripotent cell  
5 of the present invention include, but are not limited to: epidermal cells, pancreatic parenchymal cells, pancreatic duct cells, hepatic cells, blood cells, cardiac muscle cells, skeletal muscle cells, osteoblasts, skeletal myoblasts, neurons, vascular endothelial cells, pigment cells, smooth  
10 muscle cells, fat cells, bone cells, cartilage cells, and the like.

As used herein, the term "tissue" refers to an aggregate of cells having substantially the same function and/or form in a multicellular organism. "Tissue" is typically an aggregate of cells of the same origin, but may be an aggregate of cells of different origins as long as the cells have the same function and/or form. Therefore, when stem cells of the present invention are used to  
15 regenerate tissue, the tissue may be composed of an aggregate of cells of two or more different origins. Typically, a tissue constitutes a part of an organ. Animal tissues are separated into epithelial tissue, connective tissue, muscular tissue, nervous tissue, and the like, on a  
20 morphological, functional, or developmental basis. Plant tissues are roughly separated into meristematic tissue and permanent tissue, according to the developmental stage of the cells constituting the tissue. Alternatively, tissues may be separated into single tissues and composite tissues  
25 according to the type of cells constituting the tissue.  
30

Thus, tissues are separated into various categories.

As used herein, the term "stem cell" refers to a cell capable of self replication and pluripotency. Typically, stem cells can regenerate an injured tissue. Stem cells used herein may be, but are not limited to, embryonic stem (ES) cells or tissue stem cells (also called tissular stem cells, tissue-specific stem cells, or somatic stem cells). A stem cell may be an artificially produced cell (e.g., fusion cells, reprogrammed cells, or the like used herein), as long as it has the above-described abilities. Embryonic stem cells are pluripotent stem cells derived from early embryos. An embryonic stem cell was first established in 1981, and has been applied to the production of knockout mice since 1989. In 1998, a human embryonic stem cell was established, which is currently becoming available for regenerative medicine. Tissue stem cells have a relatively limited level of differentiation, unlike embryonic stem cells. Tissue stem cells are present in tissues and have an undifferentiated intracellular structure. Tissue stem cells have a higher nucleus/cytoplasm ratio and have few intracellular organelles. Most tissue stem cells have pluripotency, a long cell cycle, and proliferative ability beyond the life of the individual. As used herein, stem cells may preferably be embryonic stem cells, though tissue stem cells may also be employed, depending on the circumstance.

Tissue stem cells are separated into categories of sites from which the cells are derived, such as the dermal

system, the digestive system, the bone marrow system, the nervous system, and the like. Tissue stem cells in the dermal system include epidermal stem cells, hair follicle stem cells, and the like. Tissue stem cells in the digestive  
5 system include pancreatic (common) stem cells, liver stem cells, and the like. Tissue stem cells in the bone marrow system include hematopoietic stem cells, mesenchymal stem cells, and the like. Tissue stem cells in the nervous system include neural stem cells, retinal stem cells, and the like.

10

As used herein, the term "somatic cell" refers to any cell other than a germ cell, such as an egg, a sperm, or the like, which does not transfer its DNA to the next generation. Typically, somatic cells have limited or no  
15 pluripotency. Somatic cells used herein may be naturally-occurring or genetically modified.

The origin of a stem cell is categorized into the ectoderm, endoderm, or mesoderm. Stem cells of  
20 ectodermal origin are mostly present in the brain, including neural stem cells. Stem cells of endodermal origin are mostly present in bone marrow, including blood vessel stem cells, hematopoietic stem cells, mesenchymal stem cells, and the like. Stem cells of mesoderm origin are mostly  
25 present in organs, including liver stem cells, pancreatic stem cells, and the like. Somatic cells may be herein derived from any germ layer. Preferably, somatic cells, such as lymphocytes, spleen cells or testis-derived cells, may be used.

30

As used herein, the term "isolated" means that naturally accompanying material is at least reduced, or preferably substantially completely eliminated, in the normal environment. Therefore, the term "isolated cell" refers to a cell substantially free from other accompanying substances (e.g., other cells, proteins, nucleic acids, etc.) in the natural environment. The term "isolated" in relation to nucleic acids or polypeptides means that, for example, the nucleic acids or the polypeptides are substantially free from cellular substances or culture media when they are produced by recombinant DNA techniques; or precursory chemical substances or other chemical substances when they are chemically synthesized. Isolated nucleic acids are preferably free from sequences that naturally flank the nucleic acid within an organism from which the nucleic acid is derived (i.e., sequences positioned at the 5' terminus and the 3' terminus of the nucleic acid).

As used herein, the term "established" in relation to cells refers to a state of a cell in which a particular property (pluripotency) of the cell is maintained and the cell undergoes stable proliferation under culture conditions. Therefore, established stem cells maintain pluripotency.

25

As used herein, the term "differentiated cell" refers to a cell having a specialized function and form (e.g., muscle cells, neurons, etc.). Unlike stem cells, differentiated cells have no or little pluripotency. Examples of differentiated cells include epidermic cells,

30



pancreatic parenchymal cells, pancreatic duct cells, hepatic cells, blood cells, cardiac muscle cells, skeletal muscle cells, osteoblasts, skeletal myoblasts, neurons, vascular endothelial cells, pigment cells, smooth muscle  
5 cells, fat cells, bone cells, cartilage cells, and the like.

As used herein, the term "state" refers to a condition concerning various parameters of a cell (e.g., cell cycle, response to an external factor, signal  
10 transduction, gene expression, gene transcription, etc.). Examples of such a state include, but are not limited to, differentiated states, undifferentiated states, responses to external factors, cell cycles, growth states, and the like.

15

As used herein, the terms "differentiation" or "cell differentiation" refers to a phenomenon where two or more types of cells having qualitative differences in form and/or function occur in a daughter cell population derived  
20 from the division of a single cell. Therefore, "differentiation" includes a process during which a population (family tree) of cells, which do not originally have a specific detectable feature, acquire a feature, such as the production of a specific protein, or the like. At  
25 present, cell differentiation is generally considered to be a state of a cell in which a specific group of genes in the genome are expressed. Cell differentiation can be identified by searching for intracellular or extracellular agents or conditions which elicit the above-described state  
30 of gene expression. Differentiated cells are stable in

As used herein, the term "pluripotency" refers to a nature of a cell, i.e., an ability to differentiate into one or more, preferably two or more, tissues or organs. Therefore, the terms "pluripotent" and "undifferentiated" are herein used interchangeably unless otherwise mentioned. Typically, the pluripotency of a cell is limited during development, and in an adult, cells constituting a tissue or organ rarely differentiate into different cells, that is, the pluripotency is usually lost. Particularly, epithelial cells resist altering into other types of epithelial cells. Such alteration typically occurs in pathological conditions, and is called metaplasia. However, mesenchymal cells tend to easily undergo metaplasia, i.e., alter to other mesenchymal cells, with relatively simple stimuli. Therefore, mesenchymal cells have a high level of pluripotency. Embryonic stem cells have pluripotency. Tissue stem cells have pluripotency. Thus, the term "pluripotency" may include the concept of totipotency. An example of an *in vitro* assay for determining whether or not a cell has pluripotency, includes, but is not limited to, culturing under conditions for inducing the formation and differentiation of embryoid bodies. Examples of an *in vivo* assay for determining the presence or absence of pluripotency, include, but are not limited to, implantation of a cell into an immunodeficient mouse so as to form teratoma.

injection of a cell into a blastocyst so as to form a chimeric embryo, implantation of a cell into a tissue of an organism (e.g., injection of a cell into ascites) so as to undergo proliferation, and the like. As used herein, one type of pluripotency is "totipotency", which refers to the ability to be differentiated into all kinds of cells which constitute an organism. The idea of pluripotency encompasses totipotency. An example of a totipotent cell is a fertilized ovum. An ability to be differentiated into only one type of cell is called "unipotency".

(Biochemistry and Molecular Biology)

As used herein, the term "agent" may refer to any substance or element as long as an intended object can be achieved (e.g., energy, such as ionizing radiation, radiation, light, acoustic waves, and the like). Examples of such a substance include, but are not limited to, proteins, polypeptides, oligopeptides, peptides, polynucleotides, oligonucleotides, nucleotides, nucleic acids (e.g., DNA such as cDNA, genomic DNA and the like, or RNA such as mRNA, RNAi and the like), polysaccharides, oligosaccharides, lipids, low molecular weight organic molecules (e.g., hormones, ligands, information transduction substances, low molecular weight organic molecules, molecules synthesized by combinatorial chemistry, low molecular weight molecules usable as medicaments (e.g., low molecular weight molecule ligands, etc.), etc.), and composite molecules thereof. External agents may be used singly or in combination. Examples of an agent specific to a polynucleotide include, but are not limited to,

representatively, a polynucleotide having complementarity to the sequence of the polynucleotide with a predetermined sequence homology (e.g., 70% or more sequence identity), a polypeptide such as a transcriptional agent binding to a promoter region, and the like. Examples of an agent specific to a polypeptide include, but are not limited to, representatively, an antibody specifically directed to the polypeptide or derivatives or analogs thereof (e.g., single chain antibody), a specific ligand or receptor when the polypeptide is a receptor or ligand, a substrate when the polypeptide is an enzyme, and the like.

As used herein the term "biological agent" refers to an agent relating to a biological organism (for example, a cell). Preferably, an agent present in a cell in a normal state is referred to a biological agent. Such biological agents include, but are not limited to, for example: nucleic acid molecules, proteins, sugars, lipids, metabolites, low molecular weight molecules, and complexes thereof, and agents including time elements and the like. Alternatively, it should be understood that such biological agents include electric current, electric potential (such as membrane potential), pH, osmotic pressure and the like in the present invention. Useful biological agents as used herein include, for example, transcriptional controlling sequence (for example, promoters and the like), structural genes, and nucleic acids encoding the same. As used herein a "collection" of "biological agents" refer to a plurality of biological agents (of the same or different types). Preferably, the collection refers to biological agents which

cooperate with each other.

As used herein, the term "gene" refers to an element defining a genetic trait. A gene is typically arranged in a given sequence on a chromosome. A gene which defines the primary structure of a protein is called a structural gene. A gene which regulates the expression of a structural gene is called a regulatory gene (e.g., promoter). Genes herein include structural genes and regulatory genes unless otherwise specified. Therefore, the term "cyclin gene" typically includes the structural gene of cyclin and the promoter of cyclin. As used herein, "gene" may refer to "polynucleotide", "oligonucleotide", "nucleic acid", and "nucleic acid molecule" and/or "protein", "polypeptide", "oligopeptide" and "peptide". As used herein, "gene product" includes "polynucleotide", "oligonucleotide", "nucleic acid" and "nucleic acid molecule" and/or "protein", "polypeptide", "oligopeptide" and "peptide", which are expressed by a gene. Those skilled in the art understand what a gene product is, according to the context.

As used herein, the term "homology" in relation to a sequence (e.g., a nucleic acid sequence, an amino acid sequence, etc.) refers to the level of identity between two or more gene sequences. Therefore, the greater the homology between two given genes, the greater the identity or similarity between their sequences. Whether or not two genes have homology is determined by comparing their sequences directly or by a hybridization method under

stringent conditions. When two gene sequences are directly compared with each other, these genes have homology if the DNA sequences of the genes have representatively at least 50% identity, preferably at least 70% identity, more preferably at least 80%, 90%, 95%, 96%, 97%, 98%, or 99% identity with each other. As used herein, the term "similarity" in relation to a sequence (e.g., a nucleic acid sequence, an amino acid sequence, or the like) refers to the level of identity between two or more sequences when conservative substitution is regarded as positive (identical) in the above-described homology. Therefore, homology and similarity differ from each other in the presence of conservative substitutions. If no conservative substitutions are present, homology and similarity have the same value.

As used herein, the comparison of similarity, identity and homology of an amino acid sequence and a nucleotide sequence is calculated with FAST, a tool for sequence analysis using default parameters.

The terms "protein", "polypeptide", "oligopeptide" and "peptide" as used herein have the same meaning and refer to an amino acid polymer having any length. This polymer may be a straight, branched or cyclic chain. An amino acid may be a naturally-occurring or nonnaturally-occurring amino acid, or a variant amino acid. The term may include those assembled into a composite of a plurality of polypeptide chains. The term also includes a naturally-occurring or artificially modified amino acid

polymers. Such modification includes, for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification (e.g., conjugation with a labeling moiety).  
5 This definition encompasses a polypeptide containing at least one amino acid analog (e.g., nonnaturally-occurring amino acid, etc.), a peptide-like compound (e.g., peptoid), and other variants known in the art, for example. Gene products, such as extracellular matrix proteins (e.g.,  
10 fibronectin, etc.), are usually in the form of a polypeptide.

The terms "polynucleotide", "oligonucleotide", "nucleic acid molecule" and "nucleic acid" as used herein have the same meaning and refer to a nucleotide polymer  
15 having any length. This term also includes an "oligonucleotide derivative" or a "polynucleotide derivative". An "oligonucleotide derivative" or a "polynucleotide derivative" includes a nucleotide derivative, or refers to an oligonucleotide or a  
20 polynucleotide having different linkages between nucleotides from typical linkages, which are interchangeably used. Examples of such an oligonucleotide specifically include 2'-O-methyl-ribonucleotide, an oligonucleotide derivative in which a phosphodiester bond  
25 in an oligonucleotide is converted to a phosphorothioate bond, an oligonucleotide derivative in which a phosphodiester bond in an oligonucleotide is converted to a N3'-P5' phosphoroamidate bond, an oligonucleotide derivative in which a ribose and a phosphodiester bond in  
30 an oligonucleotide are converted to a peptide-nucleic acid

bond, an oligonucleotide derivative in which uracil in an oligonucleotide is substituted with C-5 propynyl uracil, an oligonucleotide derivative in which uracil in an oligonucleotide is substituted with C-5 thiazole uracil, 5 an oligonucleotide derivative in which cytosine in an oligonucleotide is substituted with C-5 propynyl cytosine, an oligonucleotide derivative in which cytosine in an oligonucleotide is substituted with phenoxazine-modified cytosine, an oligonucleotide derivative in which ribose in 10 DNA is substituted with 2'-O-propyl ribose, and an oligonucleotide derivative in which ribose in an oligonucleotide is substituted with 2'-methoxyethoxy ribose. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses 15 conservatively-modified variants thereof (e.g. degenerate codon substitutions) and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be produced by generating sequences in which the third position of one or more selected 20 (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzner et al., Nucleic Acid Res. 19:5081(1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); Rossolini et al., Mol. Cell. Probes 8:91-98(1994)). A gene encoding an extracellular matrix protein (e.g., 25 fibronectin, etc.) or the like is usually in the form of polynucleotide. A molecule to be transfected is in the form of a polynucleotide.

As used herein, the term "corresponding" amino 30 acid or nucleic acid refers to an amino acid or nucleotide



in a given polypeptide or polynucleotide molecule, which has, or is anticipated to have, a function similar to that of a predetermined amino acid or nucleotide in a polypeptide or polynucleotide as a reference for comparison.

5 Particularly, in the case of enzyme molecules, the term refers to an amino acid which is present at a similar position in an active site and similarly contributes to catalytic activity. For example, in the case of a transcriptional controlling sequence of a polynucleotide, it may be a portion

10 similar to that of corresponding ortholog in the particular portion of the transcription controlling sequence.

As used herein, the term "corresponding" gene (e.g., a polypeptide or polynucleotide molecule) refers to

15 a gene in a given species, which has, or is anticipated to have, a function similar to that of a predetermined gene in a species as a reference for comparison. When there are a plurality of genes having such a function, the term refers to a gene having the same evolutionary origin. Therefore,

20 a gene corresponding to a given gene may be an ortholog of the given gene. Therefore, genes corresponding to mouse cyclin genes can be found in other animals. Such a corresponding gene can be identified by techniques well known in the art. Therefore, for example, a corresponding

25 gene in a given animal can be found by searching a sequence database of the animal (e.g., human, rat) using the sequence of a reference gene (e.g., mouse cyclin gene, etc.) as a query sequence.

30 As used herein, the term "fragment" with respect

to a polypeptide or polynucleotide refers to a polypeptide or polynucleotide having a sequence length ranging from 1 to n-1 with respect to the full length of the reference polypeptide or polynucleotide (of length n). The length of the fragment can be appropriately changed depending on the purpose. For example, in the case of polypeptides, the lower limit of the length of the fragment includes 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 or more amino acids. Lengths represented by integers which are not herein specified (e.g., 11 and the like) may be appropriate as a lower limit. For example, in the case of polynucleotides, the lower limit of the length of the fragment includes 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 75, 100 or more nucleotides. Lengths represented by integers which are not herein specified (e.g., 11 and the like) may be appropriate as a lower limit. As used herein, the length of polypeptides or polynucleotides can be represented by the number of amino acids or nucleic acids, respectively. However, the above-described numbers are not absolute. The above-described numbers as the upper or lower limit are intended to include some greater or smaller numbers (e.g.,  $\pm 10\%$ ), as long as the same function is maintained. For this purpose, "about" may be herein put ahead of the numbers. However, it should be understood that the interpretation of numbers is not affected by the presence or absence of "about" in the present specification.

As used herein, the term "biological activity" refers to activity possessed by an agent (e.g., a polynucleotide, a protein, etc.) within an organism,

including activities exhibiting various functions (e.g., transcription promoting activity, etc.). For example, when a certain factor is an enzyme, the biological activity thereof includes its enzyme activity. In another example,  
5 when a certain factor is a ligand, the biological activity thereof includes the binding of the ligand to a receptor corresponding thereto. The above-described biological activity can be measured by techniques well-known in the art.

10

As used herein, the term "polynucleotides hybridizing under stringent conditions" refers to conditions commonly used and well known in the art. Such a polynucleotide can be obtained by conducting colony hybridization, plaque  
15 hybridization, Southern blot hybridization, or the like, using a polynucleotide selected from the polynucleotides of the present invention. Specifically, a filter on which DNA derived from a colony or plaque is immobilized is used to conduct hybridization at 65°C in the presence of 0.7 to  
20 1.0 M NaCl. Thereafter, a 0.1 to 2-fold concentration SSC (saline-sodium citrate) solution (1-fold concentration SSC solution is composed of 150 mM sodium chloride and 15 mM sodium citrate) is used to wash the filter at 65°C. Polynucleotides identified by this method are referred to  
25 as "polynucleotides hybridizing under stringent conditions". Hybridization can be conducted in accordance with a method described in, for example, Molecular Cloning 2nd ed., Current Protocols in Molecular Biology, Supplement 1-38, DNA Cloning 1: Core Techniques, A Practical Approach,  
30 Second Edition, Oxford University Press (1995), and the like.

Here, sequences hybridizing under stringent conditions exclude, preferably, sequences containing only A or T. "Hybridizable polynucleotide" refers to a polynucleotide which can hybridize to other polynucleotides under the  
5 above-described hybridization conditions. Specifically, the hybridizable polynucleotide includes at least a polynucleotide having a homology of at least 60% to the base sequence of DNA encoding a polypeptide having an amino acid sequence specifically herein disclosed, preferably a  
10 polynucleotide having a homology of at least 80%, and more preferably a polynucleotide having a homology of at least 95%.

As used herein, the term "salt" has the same  
15 meaning as that commonly understood by those skilled in the art, including both inorganic and organic salts. Salts are typically generated by neutralizing reactions between acids and bases. Salts include NaCl, K<sub>2</sub>SO<sub>4</sub>, and the like, which are generated by neutralization, and in addition, PbSO<sub>4</sub>,  
20 ZnCl<sub>2</sub>, and the like, which are generated by reactions between metals and acids. The latter salts may not be generated directly by neutralizing reactions, but may be regarded as a product of neutralizing reactions between acids and bases. Salts may be divided into the following categories: normal  
25 salts (salts without any H-groups from acids or without any OH-groups from bases, including, for example, NaCl, NH<sub>4</sub>Cl, CH<sub>3</sub>COONa, and Na<sub>2</sub>CO<sub>3</sub>), acid salts (salts with remaining H-groups from acids, including, for example, NaHCO<sub>3</sub>, KHSO<sub>4</sub>, and CaHPO<sub>4</sub>), and basic salts (salts with remaining OH-groups  
30 from bases, including, for example, MgCl(OH) and CuCl(OH)).

This classification is not very important in the present invention. Examples of preferable salts include salts constituting media (e.g., calcium chloride, sodium hydrogen phosphate, sodium hydrogen carbonate, sodium pyruvate, 5 HEPES, sodium chloride, potassium chloride, magnesium sulfide, iron nitrate, amino acids, vitamins, etc.), salts constituting buffer (e.g., calcium chloride, magnesium chloride, sodium hydrogen phosphate, sodium chloride, etc.), and the like. These salts are preferable as they have a high 10 affinity for cells and thus are better able to maintain cells in culture. These salts may be used singly or in combination. Preferably, these salts may be used in combination. This is because a combination of salts tends to have a higher affinity for cells. Therefore, a plurality of salts (e.g., 15 calcium chloride, magnesium chloride, sodium hydrogen phosphate, and sodium chloride) are preferably contained in a medium, rather than only NaCl or the like. More preferably, all salts suitable for cell culture medium may be added to the medium. In another preferred embodiment, 20 glucose may be added to medium.

As used herein the term "material" or "substance" is used in the broadest meaning as used in the art to refer to any thing that is positively or negatively 25 charged.

As used herein, the term "positively charged substance" encompasses all substances having a positive charge. Such substances include cationic substances such 30 as cationic polymers, cationic lipids and the like, but are

not limited to these. Advantageously, such positively charged substances can form a complex. Such positively charged substances which can form a complex include, for example, substances having a certain molecular weight (for example, cationic polymers) and substances which can remain insoluble, that is, without being dissolved to a certain extent in a specific solvent such as water, an aqueous solution or the like (for example, cationic lipids), but are not limited to these. Preferable positively charged substances include, for example, polyethylene imine, poly-L-lysine, synthetic polypeptides, or derivatives thereof, but are not limited to these. Positively charged substances include, for example, biological molecules such as histone and synthetic polypeptides, but are not limited to these. The type of preferable positively charged substances changes in accordance with the type of negatively charged substances, which act as a complex partner to form complexes with the positively charged substances. It requires no specific creativity for those skilled in the art to select a preferable complex partner using technology well known in the art. For selecting a preferable complex partner, various parameters are considered including, but not limited to, charge, molecular weight, hydrophobicity, hydrophilicity, properties of substituents, pH, temperature, salt concentration, pressure, and other physical and chemical parameters.

As used herein, the term "cationic polymer" refers to a polymer having a cationic functional group, and encompasses, for example, polyethylene imine,

poly-L-lysine, synthetic polypeptides, and derivatives thereof, but is not limited to these.

As used herein, the term "cationic lipid" refers  
5 to a lipid having a cationic functional group, and encompasses, for example, phosphatidyl choline, phosphatidyl ethanol amine, phosphatidyl serine, and derivatives thereof, but is not limited to these.

10 Cationic functional groups include, for example, primary amines, secondary amines, and tertiary amines, but are not limited to these.

As used herein, the term "negatively charged  
15 substance" encompasses all substances having a negative charge. Such substances include biological molecular polymers, anionic substances such as anionic lipids, and the like, but are not limited to these. Advantageously, such negatively charged substances can form a complex. Such  
20 negatively charged substances which can form a complex include, for example, substances having a certain molecular weight (for example, anionic polymers such as DNA) and substances which can remain insoluble, that is, without being dissolved to a certain extent in a specific solvent  
25 such as water, an aqueous solutions or the like (for example, anionic lipids), but are not limited to these. Preferable negatively charged substances include, for example, DNA, RNA, PNA, polypeptides, chemical compounds, and complexes thereof, but are not limited to these. Negatively charged  
30 substances include, for example, DNA, RNA, PNA, polypeptides,

chemical compounds, and complexes thereof, but are not limited to these. The type of preferable negatively charged substances changes in accordance with the type of positively charged substances, which act as a complex partner to form  
5 complexes with the negatively charged substances. It requires no specific creativity for those skilled in the art to select a preferable complex partner using technology well known in the art. For selecting a preferable complex partner, various parameters are considered as described  
10 above with regard to negatively charged substances.

As used herein, the term "anionic polymer" encompasses polymers having an anionic functional group, and includes, for example, DNA, RNA, PNA, polypeptides,  
15 chemical compounds, and complexes thereof, but is not limited to these.

As used herein, the term "anionic lipid" encompasses lipids having an anionic functional group, and  
20 include, for example, phosphatidic acid, phosphatidyl serine, but is not limited to these.

Anionic functional groups include, for example, carboxylic groups and phosphoric acid groups, but are not  
25 limited to these.

The type of charge of a target substance can be converted by adding a part of a substituent or the like having a positive charge or a negative charge to the target  
30 substance.



In the case where a preferable complex partner has the same type of charge as that of the target substance, formation of a complex can be promoted by converting the type of charge of either the complex partner or the target substance.

5

As used herein, the term "complex" refers to two or more substances which directly or indirectly interact with each other and as a result, act as if they were one substance as a whole.

10

As used herein, the term "complex partner" used for a certain member forming a complex refers to another member interacting with the certain member directly or indirectly.

15

As used herein, the condition for forming a complex changes in accordance with the type of complex partner. Such a condition can be easily understood by those skilled in the art. Those skilled in the art can easily form a complex from any complex partners (for example, a positively charged substance and a negatively charged substance) using a technique well known in the art.

20

As used herein, when a complex of positively and negatively charged substances is used, either or both thereof may be identical to a biological agent.

25

As used herein, the term "immobilization" used for a solid-phase support refers to a state in which a substance as a subject of immobilization (e. g., a biological

30

molecule) is held on the support for at least a certain time period, or an act-of placing the substance into such a state. As such, in the case where the condition is changed after the substance is immobilized on the solid-phase support (for  
5 example, the substance is immersed in another solvent), the substance may be released from the immobilization state.

As used herein, the term "cell affinity" refers to a property of a substance that when the substance is placed  
10 in an interactable state with a cell (e. g. germ cell, animal cell, yeast, plant cell) or an object containing a cell (e. g., tissue, organs, biological organisms), the substance does not have any adverse influence on the cell or the object containing the cell. Preferably, substances having cell  
15 affinity may be substances with which a cell interacts as a priority, but are not limited to these. According to the present invention, the substance to be immobilized (e. g., positively charged substances and/or negatively charged substances) preferably have cell affinity, but cell affinity  
20 is not absolutely necessary. It was unexpectedly found that when the substance to be immobilized has cell affinity, the cell affinity of the substance is maintained or improved when the substance is immobilized according to the present invention. In light of the past situation where a substance  
25 having cell affinity does not necessarily maintain its cell affinity when immobilized on a solid-phase support, the effect of the present invention is enormous.

As used herein, the term "probe" refers to a  
30 substance for use in searching, which is used in a biological

experiment, such as *in vitro* and/or *in vivo* screening or the like, including, but not limited to, for example, a nucleic acid molecule having a specific base sequence or a peptide containing a specific amino acid sequence.

5

Examples of a nucleic acid molecule as a common probe include one having a nucleic acid sequence having a length of at least 8 contiguous nucleotides, which is homologous or complementary to the nucleic acid sequence of a gene of interest. Such a nucleic acid sequence may be preferably a nucleic acid sequence having a length of at least 9 contiguous nucleotides, more preferably a length of at least 10 contiguous nucleotides, and even more preferably a length of at least 11 contiguous nucleotides, a length of at least 12 contiguous nucleotides, a length of at least 13 contiguous nucleotides, a length of at least 14 contiguous nucleotides, a length of at least 15 contiguous nucleotides, a length of at least 20 contiguous nucleotides, a length of at least 25 contiguous nucleotides, a length of at least 30 contiguous nucleotides, a length of at least 40 contiguous nucleotides, or a length of at least 50 contiguous nucleotides. A nucleic acid sequence used as a probe includes a nucleic acid sequence having at least 70% homology to the above-described sequence, more preferably at least 80%, and even more preferably at least 90% or at least 95%.

As used herein, the term "search" indicates that a given nucleic acid sequence is utilized to find other nucleic acid base sequences having a specific function

30

and/or property either electronically or biologically, or using other methods. Examples of an electronic search include, but are not limited to, BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)), FASTA (Pearson & Lipman, Proc. Natl. Acad. Sci., USA 85:2444-2448 (1988)), the Smith and Waterman method (Smith and Waterman, J. Mol. Biol. 147:195-197 (1981)), and the Needleman and Wunsch method (Needleman and Wunsch, J. Mol. Biol. 48:443-453 (1970)), and the like. Examples of a biological search include, but are not limited to, a macroarray in which genomic DNA is attached to a nylon membrane or the like or a microarray (microassay) in which genomic DNA is attached to a glass plate under stringent hybridization conditions, PCR, in situ hybridization, and the like.

15

As used herein, the term "primer" refers to a substance required for the initiation of a reaction of a macromolecule compound to be synthesized, in a macromolecule synthesis enzymatic reaction. In a reaction for synthesizing a nucleic acid molecule, a nucleic acid molecule (e.g., DNA, RNA, or the like) which is complementary to part of a macromolecule compound to be synthesized may be used.

20

25

A nucleic acid molecule which is ordinarily used as a primer includes one that has a nucleic acid sequence having a length of at least 8 contiguous nucleotides, which is complementary to the nucleic acid sequence of a gene of interest. Such a nucleic acid sequence preferably has a length of at least 9 contiguous nucleotides, more preferably

30

a length of at least 10 contiguous nucleotides, even more preferably a length of at least 11 contiguous nucleotides, a length of at least 12 contiguous nucleotides, a length of at least 13 contiguous nucleotides, a length of at least 14 contiguous nucleotides, a length of at least 15 contiguous nucleotides, a length of at least 16 contiguous nucleotides, a length of at least 17 contiguous nucleotides, a length of at least 18 contiguous nucleotides, a length of at least 19 contiguous nucleotides, a length of at least 20 contiguous nucleotides, a length of at least 25 contiguous nucleotides, a length of at least 30 contiguous nucleotides, a length of at least 40 contiguous nucleotides, and a length of at least 50 contiguous nucleotides. A nucleic acid sequence used as a primer includes a nucleic acid sequence having at least 70% homology to the above-described sequence, more preferably at least 80%, even more preferably at least 90%, and most preferably at least 95%. An appropriate sequence as a primer may vary depending on the property of the sequence to be synthesized (amplified). Those skilled in the art can design an appropriate primer depending on the sequence of interest. Such primer design is well known in the art and may be performed manually or using a computer program (e.g., LASERGENE, Primer Select, DNASTar).

As used herein, the term "epitope" refers to an antigenic determinant. Therefore, the term "epitope" includes a set of amino acid residues which are involved in recognition by a particular immunoglobulin, or in the context of T cells, those residues necessary for recognition by the T cell receptor proteins and/or Major

Histocompatibility Complex (MHC) receptors. This term is also used interchangeably with "antigenic determinant" or "antigenic determinant site". In the field of immunology, *in vivo* or *in vitro*, an epitope is the features of a molecule (e.g., primary, secondary and tertiary peptide structure, and charge) that form a site recognized by an immunoglobulin, T cell receptor or HLA molecule. An epitope including a peptide comprises 3 or more amino acids in a spatial conformation which is unique to the epitope. Generally, an epitope consists of at least 5 such amino acids, and more ordinarily, consists of at least 6, 7, 8, 9 or 10 such amino acids. The greater the length of an epitope, the more the similarity of the epitope to the original peptide, i.e., longer epitopes are generally preferable. This is not necessarily the case when the conformation is taken into account. Methods of determining the spatial conformation of amino acids are known in the art, and include, for example, X-ray crystallography and 2-dimensional nuclear magnetic resonance spectroscopy. Furthermore, the identification of epitopes in a given protein is readily accomplished using techniques well known in the art. See, also, Geysen et al., Proc. Natl. Acad. Sci. USA (1984) 81: 3998 (general method of rapidly synthesizing peptides to determine the location of immunogenic epitopes in a given antigen); U. S. Patent No. 4,708,871 (procedures for identifying and chemically synthesizing epitopes of antigens); and Geysen et al., Molecular immunology (1986) 23: 709 (technique for identifying peptides with high affinity for a given antibody). Antibodies that recognize the same epitope can be identified in a simple immunoassay. Thus, methods for

determining epitopes including a peptide are well known in the art. Such an epitope can be determined using a well-known, common technique by those skilled in the art if the primary nucleic acid or amino acid sequence of the epitope is provided.

Therefore, an epitope including a peptide requires a sequence having a length of at least 3 amino acids, preferably at least 4 amino acids, more preferably at least 5 amino acids, at least 6 amino acids, at least 7 amino acids, at least 8 amino acids, at least 9 amino acids, at least 10 amino acids, at least 15 amino acids, at least 20 amino acids, and 25 amino acids. Epitopes may be linear or conformational.

As used herein, the term "agent binding specifically to" a certain nucleic acid molecule or polypeptide refers to an agent which has a level of binding to the nucleic acid molecule or polypeptide equal to or higher than a level of binding to other nucleic acid molecules or polypeptides. Examples of such an agent include, but are not limited to, when a target is a nucleic acid molecule, a nucleic acid molecule having a complementary sequence of a nucleic acid molecule of interest, a polypeptide capable of binding to a nucleic acid sequence of interest (e.g., a transcription agent, etc.), and the like, and when a target is a polypeptide, an antibody, a single chain antibody, either of a pair of a receptor and a ligand, either of a pair of an enzyme and a substrate, and the like.

As used herein, the term "antibody" encompasses polyclonal antibodies, monoclonal antibodies, human antibodies, humanized antibodies, polyfunctional antibodies, chimeric antibodies, and anti-idiotypic antibodies, and fragments thereof (e.g., F(ab')<sub>2</sub> and Fab fragments), and other recombinant conjugates. These antibodies may be fused with an enzyme (e.g., alkaline phosphatase, horseradish peroxidase,  $\alpha$ -galactosidase, and the like) via a covalent bond or by recombination.

As used herein, the term "monoclonal antibody" refers to an antibody composition having a group of homologous antibodies. This term is not limited by the production manner thereof. This term encompasses all immunoglobulin molecules and Fab molecules, F(ab')<sub>2</sub> fragments, Fv fragments, and other molecules having an immunological binding property of the original monoclonal antibody molecule. Methods for producing polyclonal antibodies and monoclonal antibodies are well known in the art, and will be more sufficiently described below.

Monoclonal antibodies are prepared by using the standard technique well known in the art (e.g., Kohler and Milstein, Nature (1975) 256:495) or a modification thereof (e.g., Buck et al. (1982) In Vitro 18:377). Representatively, a mouse or rat is immunized with a protein bound to a protein carrier, and boosted. Subsequently, the spleen (and optionally several large lymph nodes) is removed and dissociated into a single cell suspension. If desired,



the spleen cells may be screened (after removal of nonspecifically adherent cells) by applying the cell suspension to a plate or well coated with a protein antigen. B-cells that express membrane-bound immunoglobulin specific for the antigen bind to the plate, and are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas. The hybridomas are used to produce monoclonal antibodies.

10

As used herein, the term "antigen" refers to any substrate to which an antibody molecule may specifically bind. As used herein, the term "immunogen" refers to an antigen capable of initiating activation of the antigen-specific immune response of a lymphocyte.

15

In a given protein molecule, a given amino acid may be substituted with another amino acid in a structurally important region, such as a cationic region or a substrate molecule binding site, without a clear reduction or loss of interactive binding ability. A given biological function of a protein is defined by the interactive ability or other property of the protein. Therefore, a particular amino acid substitution may be performed in an amino acid sequence, or at the DNA sequence level, to produce a protein which maintains the original property after the substitution. Therefore, various modifications of peptides as disclosed herein and DNA encoding such peptides may be performed without clear losses of biological activity.

25  
30

When the above-described modifications are designed, the hydrophobicity indices of amino acids may be taken into consideration. The hydrophobic amino acid indices play an important role in providing a protein with an interactive biological function, which is generally recognized in the art (Kyte, J. and Doolittle, R.F., J. Mol. Biol. 157(1):105-132, 1982). The hydrophobic property of an amino acid contributes to the secondary structure of a protein and then regulates interactions between the protein and other molecules (e.g., enzymes, substrates, receptors, DNA, antibodies, antigens, etc.). Each amino acid is given a hydrophobicity index based on the hydrophobicity and charge properties thereof, as follows: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamic acid (-3.5); glutamine (-3.5); aspartic acid (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is well known that if a given amino acid is substituted with another amino acid having a similar hydrophobicity index, the resultant protein may still have a biological function similar to that of the original protein (e.g., a protein having an equivalent enzymatic activity). For such an amino acid substitution, the hydrophobicity index is preferably within  $\pm 2$ , more preferably within  $\pm 1$ , and even more preferably within  $\pm 0.5$ . It is understood in the art that such an amino acid substitution based on

hydrophobicity is efficient.

Hydrophilicity may also be considered for conservative substitution. As described in US Patent No. 4,554,101, amino acid residues are given the following hydrophilicity indices: arginine (+3.0); lysine (+3.0); aspartic acid (+3.0 $\pm$ 1); glutamic acid (+3.0 $\pm$ 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 $\pm$ 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4). It is understood that an amino acid may be substituted with another amino acid which has a similar hydrophilicity index and can still provide a biological equivalent. For such an amino acid substitution, the hydrophilicity index is preferably within  $\pm 2$ , more preferably  $\pm 1$ , and even more preferably  $\pm 0.5$ .

(Profile and its relevant techniques)

As used herein, the term "profile" in relation to a cell refers to a set of measurements of the biological state of the cell. Particularly, the term "profile of a cell" refers to a set of discrete or continuous values obtained by quantitatively measuring a level of a "cellular component". A level of a cellular component includes the expression level of a gene, the transcription level of a gene (the activity level of a transcription control sequence), the amount of mRNA encoding a specific gene, and the expression level of a protein in biological systems. The level of each cellular component, such as the expression

level of mRNA and/or protein, is known to alter in response to treatment with drugs or cellular biological perturbation or vibration. Therefore, the measurement of a plurality of "cellular components" generates a large amount of  
5 information about the effects of stimuli on the biological state of a cell. Therefore, the profile is more and more important in the analysis of cells. Mammalian cells contain about 30,000 or more cellular components. Therefore, the profile of an individual cell is usually complicated. A  
10 profile in a predetermined state of a biological system may often be measured after stimulating the biological system. Such stimulation is performed under experimental or environmental conditions associated with the biological system. Examples of a stimulus include exposure of a  
15 biological system to a drug candidate, introduction of an exogenous gene, passage of time, deletion of a gene from the system, alteration of culture conditions, and the like. The wide range measurement of cellular components (i.e., profiles of gene replication or transcription, protein  
20 expression, and response to stimuli) has a high level of utility including comparison and investigation of the effects of drugs, diagnosis of diseases, and optimization of drug administration to patients as well as investigation of cells. Further, profiles are useful for basic life  
25 science research. Such profile data may be produced and presented as data in a variety of formats. Such formats include, but are not limited to: a function between a numerical value and a period of time, a graphic format, a image format and the like. Accordingly, data relating to  
30 a profile may also be called "profile data" as used herein.

Such data production may readily be carried out by using a computer. Coding of an appropriate program may also be carried out by using well known technology in the art.

5                   In the cell analysis of the present application, as regards to information derived from a cell or a substance interacting with the cell, a variety of processes and means for detection may be used. Such process and means for detection include, but are not limited to: those using visual  
10 inspection, optical microscopes, fluorescence microscopes, reading apparatus using a laser light source, surface plasmon resonance (SPR) imaging, electric signal, chemical and biochemical markers, or a combination thereof.

15                   As used herein, the term "time-lapse profile" in relation to a certain cell refers to a profile which indicates time-lapse changes in a parameter relating to the cell. Examples of a time-lapse profile include, but are not limited to, a time-lapse profile of transcription levels,  
20 a time-lapse profile of expression levels (translation levels), a time-lapse profile of signal transduction, a time-lapse profile of neural potential, and the like. A time-lapse profile may be produced by continuously recording a certain parameter (e.g., a signal caused by a label  
25 associated with a transcription level). Time-lapse measurement may mean continuous measurement. Therefore, the term "time-lapse profile" as used herein may also be referred to as "continuous profile".

30                   As used herein the term "information" of a cell

refers to those acting to direct an object as a whole by binding a number of elements present in the cell. A collection of information can be said to constitute a digital cell.

5

As used herein, the term "state" refers to a condition concerning various parameters of a cell (e.g., cell cycle, response to an external factor, signal transduction, gene expression, gene transcription, etc.).

10 Examples of such a state include, but are not limited to, differentiated states, undifferentiated states, responses to external factors, cell cycles, growth states, and the like. The responsiveness or resistance of an organism of interest with respect to the following parameters of the,

15 particularly, environment of the organism may be used herein as a measure of the state of the organism: temperature, humidity (e.g., absolute humidity, relative humidity, etc.), pH, salt concentration (e.g., the concentration of all salts or a particular salt), nutrients (e.g., the amount of

20 carbohydrate, etc.), metals (e.g., the amount or concentration of all metals or a particular metal (e.g., a heavy metal, etc.)), gas (e.g., the amount of all gases or a particular gas), organic solvent (e.g., the amount of all organic solvents or a particular organic solvent (e.g.,

25 ethanol, etc.)), pressure (e.g., local or global pressure, etc.), atmospheric pressure, viscosity, flow rate (e.g., the flow rate of a medium in which an organism is present, etc.), light intensity (e.g., the quantity of light having a particular wavelength, etc.), light wavelength (e.g.,

30 visible light, ultraviolet light, infrared light, etc.),

electromagnetic waves, radiation, gravity, tension,  
acoustic waves, organisms other than an organism of interest  
(e.g., parasites, pathogenic bacteria, etc.), chemicals  
(e.g., pharmaceuticals, etc.), antibiotics,  
5 naturally-occurring substances, metal stresses, physical  
stresses, and the like.

As used herein, the term "environment" (or  
"Umgebung" in German) in relation to an entity refers to  
10 a circumstance which surrounds the entity. In an  
environment, various components and quantities of state are  
recognized, which are called environmental factors.  
Examples of environmental factors include the  
above-described parameters. Environmental factors are  
15 typically roughly divided into non-biological  
environmental factors and biological environmental factors.  
Non-biological environmental factors (inorganic  
environment factors) may be divided into physical factors  
and chemical factors, or alternatively, climatic factors  
20 and soil factors. Various environmental factors do not  
always act on organisms independently, but may be associated  
with one another. Therefore, environment factors may be  
herein observed one by one or as a whole (a whole of various  
parameters). It has been believed that it was difficult to  
25 maintain such an environment in a consistent state. This  
is particularly the case since it has been difficult to  
maintain cells and to immobilize cells, and to introduce  
substances such as nucleic acids into a cell. The present  
invention has also solved at least one of these problems.  
30 As used herein the term "consistent environment" refers to

substantially all of the circumstances surrounding a cell of interest. Accordingly, as long as a cell can grow or differentiate in a similar manner, such environments are deemed to be consistent environments. As used herein, a  
5 consistent environment refers to an environment where the parameters are the same except for a specific stimulus (for example, an external stimulus).

Examples of such an environment include at least  
10 one factor, as a parameter, selected from the group consisting of temperature, humidity, pH, salt concentration, nutrients, metal, gas, organic solvent, pressure, atmospheric pressure, viscosity, flow rate, light intensity, light wavelength, electromagnetic waves, radiation,  
15 gravity, tension, acoustic waves, organisms (e.g., parasites, etc.) other than the organism, chemical agents, antibiotics, natural substances, mental stress, and physical stress, and any combination thereof.

20 Examples of temperature include, but are not limited to, high temperature, low temperature, very high temperature (e.g., 95°C, etc.), very low temperature (e.g., -80°C, etc.), a wide range of temperature (e.g., 150 to -270°C, etc.), and the like.

25 Examples of humidity include, but are not limited to, a relative humidity of 100%, a relative humidity of 0%, an arbitrary point from 0% to 100%, and the like.

30 Examples of pH include, but are not limited to,



an arbitrary point from 0 to 14, and the like.

Examples of salt concentration include, but are not limited to, a NaCl concentration (e.g., 3%, etc.), an  
5 arbitrary point of other salt concentrations from 0 to 100%, and the like.

Examples of nutrients include, but are not limited to, proteins, glucose, lipids, vitamins, inorganic  
10 salts, and the like.

Examples of metals include, but are not limited to, heavy metals (e.g., mercury, cadmium, etc.), lead, gold, uranium, silver, and the like.  
15

Examples of gas include, but are not limited to, oxygen, nitrogen, carbon dioxide, carbon monoxide, and a mixture thereof, and the like.

20 Examples of organic solvents include, but are not limited to, ethanol, methanol, xylene, propanol, and the like.

Examples of pressure include, but are not  
25 limited to, an arbitrary point from 0 to 10 ton/cm<sup>2</sup>, and the like.

Examples of atmospheric pressure include, but are not limited to, an arbitrary point from 0 to 100  
30 atmospheric pressure, and the like.

Examples of viscosity include, but are not limited to the viscosity of any fluid (e.g., water, glycerol, etc.) or a mixture thereof, and the like.

5

Examples of flow rate include, but are not limited to an arbitrary point from 0 to the velocity of light.

Examples of light intensity include, but are not limited to, a point between darkness and the level of sunlight.

Examples of light wavelength include, but are not limited to visible light, ultraviolet light (UV-A, UV-B, UV-C, etc.), infrared light (far infrared light, near infrared light, etc.), and the like.

Examples of electromagnetic waves include ones having an arbitrary wavelength.

20

Examples of radiation include ones having an arbitrary intensity.

Examples of gravity include, but are not limited to, an arbitrary gravity on the Earth or an arbitrary point from zero gravity to the gravity on the Earth, or an arbitrary gravity greater than or equal to a gravity on the Earth.

Examples of tension include ones having an arbitrary strength.

30

Examples of acoustic waves include ones having an arbitrary intensity and wavelength.

5                Examples of organisms other than an organism of interest include, but are not limited to, parasites, pathogenic bacteria, insects, nematodes, and the like.

                 Examples of chemicals include, but are not  
10 limited to hydrochloric acid, sulfuric acid, sodium hydroxide, and the like.

                 Examples of antibiotics include, but are not limited to, penicillin, kanamycin, streptomycin, quinoline,  
15 and the like.

                 Examples of naturally-occurring substances include, but are not limited to, puffer toxin, snake venom, alkaloid, and the like.  
20

                 Examples of mental stress include, but are not limited to starvation, population density, confined spaces, high places, and the like.

25                Examples of physical stress include, but are not limited to vibration, noise, electricity, impact, and the like.

                 As used herein when referring to a digital cell of the  
30 present invention, the environment is presented as an

"environment parameter". Such environment parameters include, but are not limited to, medium (type, composition), pH, temperature, moisture, CO<sub>2</sub> concentration, O<sub>2</sub> concentration, the presence or absence of an antibiotic, 5 the presence or absence of a particular nutrient and the like.

As used herein the term "stimulant" refers to an acting agent which causes or induces expression or 10 enhancement of a specific living action given to a cell from outside. Stimuli include, but are not limited to: a physical stimulus, a chemical stimulus, a biological stimulus, a biochemical stimulus, and the like. Physical stimuli include, but are not limited to: for example, light, 15 electric waves, electric current, pressure, sound (vibration) and the like. Chemical stimuli include but are not limited to: for example, stimuli from chemicals such as antibiotics, nutrients, vitamins, metals, ions, acids, alkalis, salts, buffers and the like. Biological stimuli 20 include, but are not limited to: for example, the existence of another organism such as the existence of a parasitic organism or the density of a cell population and the like. Biochemical stimuli include, but are not limited to the existence of cell signaling transduction agents, and the 25 like.

As used herein, when the digital cell of the present invention is used, a stimulus is presented as a "stimulus parameter". A stimulus parameter corresponding to those in 30 response to any stimulus as described herein may be used.

As used herein, it should be understood that the stimulus parameter includes agents for transducing a stimulus such as a reporter and the like. Such reporters include, but are not limited to: for example, on-off regulation of expression  
5 against an antibiotic, a transcription-controlling sequence, radioactivity, fluorophores and the like.

As used herein the term "response" to a stimulus refers to any response of a cell to a stimulus such as a  
10 change in cell morphology, change in metabolism, change in other cellular behaviors, change in signal transduction and the like. Therefore, for example, results of experiments using the digital cell of the present invention may be recorded as cell dynamics data. Alternatively, when using  
15 the above reporter, the result of such a response to the stimulus may be raw data of the reporter, or data transformed from the data of the reporter.

As used herein, the term "transcription control  
20 sequence" refers to a sequence which can regulate the transcription level of a gene. Such a sequence is at least two nucleotides in length. Examples of such a sequence include, but are not limited to, promoters, enhancers, silencers, terminators, sequences flanking other genomic  
25 structural genes, genomic sequences other than exons, sequences within exons, and the like. A transcription control sequence used herein is not related to a particular type. Rather, important information about a transcription control sequence is time-lapse fluctuation. Such  
30 fluctuation is referred to as a process (changes in a state

of a cell). Therefore, such a transcription control sequence may be herein arbitrarily selected. Such a transcription control sequence may include those which are not conventionally used as markers. Preferably, a  
5 transcription control sequence has the ability to bind to a transcription factor.

As used herein, the term "transcription factor" refers to a factor which regulates the process of  
10 transcription of a gene. The term "transcription factor" mainly indicates a factor which regulates a transcription initiation reaction. Transcription factors are roughly divided into the following groups: basic transcription factors required for placing an RNA polymerase into a  
15 promoter region on DNA; and transcription regulatory factors which bind to cis-acting elements present upstream or downstream of a transcription region to regulate the synthesis initiation frequency of RNA.

20 Basic transcription factors are prepared depending on the type of RNA polymerase. A TATA-binding protein is believed to be common to all transcription systems. Although there are a number of types of transcription factors, a typical transcription factor consists of a portion  
25 structurally required for binding to DNA and a portion required for activating or suppressing transcription. Factors which have a DNA-binding portion and can bind to cis-acting elements are collectively referred to as trans-acting factors.

30

A portion required for activating or suppressing transcription is involved in interaction with other transcription factors or basic transcription factors. Such a portion is believed to play a role in regulating transcription via a structural change in DNA or a transcription initiating complex. Transcription regulatory factors are divided into several groups or families according to the structural properties of these portions, including factors which play an important role in the development or differentiation of a cell.

Examples of such a transcription factor include, but are not limited to, STAT1, STAT2, STAT3, GAS, NFAT, Myc, AP1, CREB, NFκB, E2F, Rb, p53, RUNX1, RUNX2, RUNX3, Nkx-2, CF2-II, Skn-1, SRY, HFH-2, Oct-1, Oct-3, Sox-5, HNF-3b, PPARγ, and the like.

As used herein, the term "terminator" refers to a sequence which is located downstream of a protein-encoding region of a gene and which is involved in the termination of transcription when DNA is transcribed into mRNA, and the addition of a poly-A sequence. It is known that a terminator contributes to the stability of mRNA, and has an influence on the level of gene expression.

25

As used herein, the term "promoter" refers to a base sequence which determines the initiation site of transcription of a gene and is a DNA region which directly regulates the frequency of transcription. Transcription is started by RNA polymerase binding to a promoter. A promoter

30

region is usually located within about 2 kbp upstream of the first exon of a putative protein coding region. Therefore, it is possible to estimate a promoter region by predicting a protein coding region in a genomic base sequence  
5 using DNA analysis software. A putative promoter region is usually located upstream of a structural gene, but is dependent on the structural gene, i.e., a putative promoter region may be located downstream of a structural gene. Preferably, a putative promoter region is located within  
10 about 2 kbp upstream of the translation initiation site of the first exon. Such promoters include, but are not limited to constitutive promoters, specific promoters and inductive promoters and the like.

15               As used herein, the term "enhancer" refers to a sequence which is used so as to enhance the expression efficiency of a gene of interest. One or more enhancers may be used, or no enhancer may be used.

20               As used herein, the term "silencer" refers to a sequence which has a function of suppressing and arresting the expression of a gene. Any silencer which has such a function may be herein used. No silencer may be used.

25               As used herein, the term "operably linked" indicates that a desired sequence is located such that expression (operation) thereof is under control of a transcription and translation regulatory sequence (e.g., a promoter, an enhancer, and the like) or a translation  
30 regulatory sequence. In order for a promoter to be operably



linked to a gene, typically, the promoter is located immediately upstream of the gene. A promoter is not necessarily adjacent to a structural gene.

5                    Sequences flanking other genome structural genes, genomic sequences other than exons, and sequences within exons may also be herein used. For example, in addition to the above-described sequences having specific names, structural gene-flanking sequences are thought to  
10 be involved in the control of transcription in terms of "processes". Therefore, such flanking sequences are also included in transcription control sequences. Genomic sequences other than exons and sequences within exons are also expected to be involved in the control of transcription  
15 in terms of "processes". Therefore, genomic sequences other than exons and sequences within exons are also included in transcription control sequences.

As used herein, the term "RNAi" is an  
20 abbreviation of RNA interference and refers to a phenomenon where an agent for causing RNAi, such as double-stranded RNA (also called dsRNA), is introduced into cells and mRNA homologous thereto is specifically degraded, so that the synthesis of gene products is suppressed, and techniques  
25 using the phenomenon. As used herein, RNAi may have the same meaning as that of an agent which causes RNAi.

As used herein, the term "an agent causing RNAi" refers to any agent capable of causing RNAi. As used herein,  
30 "an agent causing RNAi of a gene" indicates that the agent

causes RNAi relating to the gene and that the effect of RNAi is achieved (e.g., suppression of expression of the gene, and the like). Examples of such an agent causing RNAi include, but are not limited to, a sequence having at least about 70% homology to the nucleic acid sequence of a target gene or a sequence hybridizable thereto under stringent conditions, RNA containing a double-stranded portion having a length of at least 10 nucleotides or variants thereof. Here, this agent may be preferably DNA containing a 3' protruding end, and more preferably the 3' protruding end has a length of 2 or more nucleotides (e.g., 2-4 nucleotides in length).

Though not wishing to be bound by any theory, a mechanism which causes RNAi is considered to be as follows. When a molecule which causes RNAi, such as dsRNA, is introduced into a cell, an RNaseIII-like nuclease having a helicase domain (called dicer) cleaves the molecule at about 20 base pair intervals from the 3' terminus in the presence of ATP in the case where the RNA is relatively long (e.g., 40 or more base pairs). As used herein, the term "siRNA" is an abbreviation of short interfering RNA and refers to short double-stranded RNA of 10 or more base pairs which are artificially chemically synthesized or biochemically synthesized, synthesized by an organism, or produced by double-stranded RNA of about 40 or more base pairs being degraded within the organism. siRNA typically has a structure comprising 5'-phosphate and 3'-OH, where the 3' terminus projects by about 2 bases. A specific protein is bound to siRNA to form RISC

(RNA-induced-silencing-complex). This complex recognizes and binds to mRNA having the same sequence as that of siRNA and cleaves mRNA at the middle of siRNA due to RNaseIII-like enzymatic activity. It is preferable that the relationship  
5 between the sequence of siRNA and the sequence of mRNA to be cleaved as a target is a 100% match. However, base mutations at a site away from the middle of siRNA do not completely remove the cleavage activity by RNAi, leaving partial activity, while base mutations in the middle of siRNA  
10 have a large influence and the mRNA cleavage activity by RNAi is considerably lowered. By utilizing such a nature, only mRNA having a mutation can be specifically degraded. Specifically, siRNA in which the mutation is provided in the middle thereof is synthesized and is introduced into  
15 a cell. Therefore, in the present invention, siRNA per se, as well as an agent capable of producing siRNA (e.g., representatively dsRNA of about 40 or more base pairs) can be used as an agent capable of eliciting RNAi.

20               Also, though not wishing to be bound by any theory, apart from the above-described pathway, the antisense strand of siRNA binds to mRNA and siRNA functions as a primer for RNA-dependent RNA polymerase (RdRP), so that dsRNA is synthesized. This dsRNA is a substrate for a dicer  
25 again, leading to production of new siRNA. It is intended that such a reaction is amplified. Therefore, in the present invention, siRNA per se, as well as an agent capable of producing siRNA are useful. In fact, in insects and the like, for example, 35 dsRNA molecules can substantially  
30 completely degrade 1,000 or more copies of intracellular

mRNA, and therefore, it will be understood that siRNA per se, as well as an agent capable of producing siRNA, is useful.

In the present invention, double-stranded RNA  
5 having a length of about 20 bases (e.g., representatively about 21 to 23 bases) or less than about 20 bases, called siRNA, can be used. Expression of siRNA in cells can suppress expression of a pathogenic gene targeted by the siRNA. Therefore, siRNA can be used for the treatment,  
10 prophylaxis, prognosis, and the like of diseases.

The siRNA of the present invention may be in any form as long as it can elicit RNAi.

15 In another embodiment, an agent capable of causing RNAi may have a short hairpin structure having a sticky portion at the 3' terminus (shRNA; short hairpin RNA). As used herein, the term "shRNA" refers to a molecule of about 20 or more base pairs in which a single-stranded RNA  
20 partially contains a palindromic base sequence and forms a double-strand structure therein (i.e., a hairpin structure). shRNA can be artificially chemically synthesized. Alternatively, shRNA can be produced by linking sense and antisense strands of a DNA sequence in  
25 reverse directions and synthesizing RNA *in vitro* with T7 RNA polymerase using the DNA as a template. Though not wishing to be bound by any theory, it should be understood that after shRNA is introduced into a cell, the shRNA is degraded in the cell to a length of about 20 bases (e.g.,  
30 representatively 21, 22, 23 bases), and causes RNAi as with

siRNA, leading to the treatment effects of the present invention. It should be understood that such an effect is exhibited in a wide range of organisms, such as insects, plants, animals (including mammals), and the like. Thus, 5 shRNA elicits RNAi as with siRNA and therefore can be used as an effective component of the present invention. shRNA may preferably have a 3' protruding end. The length of the double-stranded portion is not particularly limited, but is preferably about 10 or more nucleotides, and more 10 preferably about 20 or more nucleotides. Here, the 3' protruding end may be preferably DNA, more preferably DNA of at least 2 nucleotides in length, and even more preferably DNA of 2-4 nucleotides in length.

15 An agent capable of causing RNAi used in the present invention may be artificially synthesized (chemically or biochemically) or naturally occurring. There is substantially no difference there between in terms of the effect of the present invention. A chemically 20 synthesized agent is preferably purified by liquid chromatography or the like.

An agent capable of causing RNAi used in the present invention can be produced *in vitro*. In this 25 synthesis system, T7 RNA polymerase and T7 promoter are used to synthesize antisense and sense RNAs from template DNA. These RNAs are annealed and thereafter introduced into a cell. In this case, RNAi is caused via the above-described mechanism, thereby achieving the effect of the present 30 invention. Here, for example, the introduction of RNA into

cell can be carried out using a calcium phosphate method.

Another example of an agent capable of causing RNAi according to the present invention is a single-stranded  
5 nucleic acid hybridizable to mRNA, or all nucleic acid  
analogs thereof. Such agents are useful for the method and  
composition of the present invention.

As used herein, the term "time-lapse" means any  
10 action or phenomenon that is related to the passage of time.

As used herein, the term "monitor" refers to the  
measurement of a state of a cell using at least one parameter  
as a measure (e.g., a labeling signal attributed to  
15 transcription, etc.). Preferably, monitoring is performed  
using a device, such as a detector, a measuring instrument,  
or the like. More preferably, such a device is connected  
to a computer for recording and/or processing data.  
Monitoring may comprise the step of obtaining image data  
20 of a solid phase support (e.g., an array, a plate, etc.).

As used herein, the term "real time" means that  
a certain state is substantially simultaneously displayed  
in another form (e.g., as an image on a display or a graph  
25 with processed data). In such a case, the "real time" lags  
behind an actual event by the time required for data  
processing. Such a time lag is included in the scope of "real  
time" if it is substantially negligible. Such a time lag  
may be typically within 10 seconds, and preferably within  
30 1 second, without limitation. A time lag exceeding

10 seconds may be included in the scope of "real time".

As used herein, the determination of a state of a cell can be performed using various methods. Examples of  
5 such methods include, but are not limited to, mathematical processing (e.g., signal processing, multivariate analysis, etc.), empirical processing, phase changes, and the like.

As used herein, the term "difference" refers to  
10 a result of mathematical processing in which a value of a control profile (e.g., without a stimulus) is subtracted from a certain profile.

As used herein, the term "phase" in relation to  
15 a time-lapse profile refers to a result of a determination of whether the profile is positive or negative with respect to a reference point (typically 0), which is expressed with + or -, and also refers to analysis based on such a result.

As used herein, the term "correlate" or  
20 "correlation" in relation to a profile (e.g., a time-lapse profile, etc.) and a state of a cell refers to an act of associating the profile or particular information about changes, with the state of the cell. A relationship between  
25 them is referred to as "correlation" or a "correlation relationship". Conventionally, it was substantially impossible to associate a profile (e.g., a time-lapse profile, etc.) with a state of a cell. No relationship between them was known. The present invention has an  
30 advantageous effect of performing such a correlation.

As used herein, correlation can be performed by associating at least one profile (e.g., a time-lapse profile, etc.) or changes therein, with a state of a cell, a tissue, an organ or an organism (e.g., drug resistance, etc.). For example, a profile (e.g., a time-lapse profile, etc.) or changes therein is quantitatively or qualitatively associated with at least one parameter indicating a state of a cell. A small number of profiles (e.g., time-lapse profile, etc.) may be used for correlation as long as correlation can be performed, typically including, without limitation, 1, preferably 2, and more preferably 3. The present invention demonstrated that at least 2, preferably at least 3, profiles (e.g., a time-lapse profile, etc.) are sufficient for specifying substantially all cells. Such an effect could not be expected by conventional profiling or assays which use point observation, and can be said to be realized by the present invention. At least one profile (e.g., a time-lapse profile, etc.) may be subjected to mathematical processing by utilizing a matrix to associate the profile with a state of a cell. In one preferred embodiment, at least 8 profiles (e.g., a time-lapse profile, etc.) may be advantageously used. By observing increases or decreases in 8 profiles, 256 results can be theoretically obtained, based on which about 300 types of cells constituting an organism can be substantially distinguished from one another. In this context, it may be more advantageous to use at least 9 or 10 structures as profiles. On the other hand, by using the technology of the present invention, it is possible to substantially understand the



state of a cell, merely by selecting any single biological agent and obtaining the profile data thereof.

Examples of a specific method for correlation  
5 include, but are not limited to, signal processing (e.g., wavelet analysis, etc.), multivariate analysis (e.g., cluster analysis, etc.), and the like.

Correlation may be performed in advance or may  
10 be performed at the time of determination of cells using a control.

As used herein, the term "external factor" in relation to a cell refers to a factor which is not usually  
15 present in the cell (e.g., a substance, energy, etc.). As used herein, the term "factor" may refer to any substance or element as long as an intended object can be achieved (e.g., energy, such as ionizing radiation, radiation, light, acoustic waves, and the like). Examples of such a substance  
20 include, but are not limited to, proteins, polypeptides, oligopeptides, peptides, polynucleotides, oligonucleotides, nucleotides, nucleic acids (e.g., DNA such as cDNA, genomic DNA and the like, or RNA such as mRNA, RNAi and the like), polysaccharides, oligosaccharides,  
25 lipids, low molecular weight organic molecules (e.g., hormones, ligands, information transduction substances, low molecular weight organic molecules, molecules synthesized by combinatorial chemistry, low molecular weight molecules usable as medicaments (e.g., low molecular  
30 weight molecule ligands, etc.), etc.), and composite

molecules thereof. External factors may be used singly or in combination. Examples of an external factor as used herein include, but are not limited to, temperature changes, humidity changes, electromagnetic wave, potential  
5 difference, visible light, infrared light, ultraviolet light, X-rays, chemical substances, pressure, gravity changes, gas partial pressure, osmotic pressure, and the like. In one embodiment, an external factor may be a biological molecule or a chemically synthesized substance.

10

As used herein, the term "biological molecule" refers to molecules relating to an organism and aggregations thereof. As used herein, the term "biological" or "organism" refers to a biological organism, including, but  
15 being not limited to, an animal, a plant, a fungus, a virus, and the like. Biological molecules include molecules extracted from an organism and aggregations thereof, though the present invention is not limited to this. Any molecule capable of affecting an organism and aggregations thereof  
20 fall within the definition of a biological molecule. Therefore, low molecular weight molecules (e.g., low molecular weight molecule ligands, etc.), capable of being used as medicaments fall within the definition of a biological molecule as long as an effect on an organism is  
25 intended. Examples of such a biological molecule include, but are not limited to, proteins, polypeptides, oligopeptides, peptides, polynucleotides, oligonucleotides, nucleotides, nucleic acids (e.g., DNA such as cDNA and genomic DNA; RNA such as mRNA),  
30 polysaccharides, oligosaccharides, lipids, low molecular

weight molecules (e.g., hormones, ligands, information transmitting substances, low molecular weight organic molecules, etc.), and composite molecules thereof and aggregations thereof (e.g., glycolipids, glycoproteins, lipoproteins, etc.), and the like. A biological molecule may include a cell itself or a portion of tissue as long as it is intended to be introduced into a cell. Typically, a biological molecule may be a nucleic acid, a protein, a lipid, a sugar, a proteolipid, a lipoprotein, a glycoprotein, a proteoglycan, or the like. Preferably, a biological molecule may include a nucleic acid (DNA or RNA) or a protein. In another preferred embodiment, a biological molecule is a nucleic acid (e.g., genomic DNA or cDNA, or DNA synthesized by PCR or the like). In another preferred embodiment, a biological molecule may be a protein. Preferably, such a biological molecule may be a hormone or a cytokine.

As used herein, the term "chemically synthesized substance" or "chemical" refers to any substance which may be synthesized by using typical chemical techniques. Such synthesis techniques are well known in the art. Those skilled in the art can produce chemically synthesized substances by combining such techniques as appropriate.

25

The term "cytokine" is used herein in the broadest sense in the art and refers to a physiologically active substance which is produced by a cell and acts on the same or a different cell. Cytokines are generally proteins or polypeptides having a function of controlling

30

an immune response, regulating the endocrine system, regulating the nervous system, acting against a tumor, acting against a virus, regulating cell growth, regulating cell differentiation, or the like. Cytokines are used  
5 herein in the form of a protein or a nucleic acid or in other forms. In actual practice, cytokines are typically proteins. The terms "growth factor" refers to a substance which promotes or controls cell growth. Growth factors are also called "proliferation factors" or "development  
10 factors". Growth factors may be added to cell or tissue culture medium, substituting for serum macromolecules. It has been revealed that a number of growth factors have a function of controlling differentiation in addition to a function of promoting cell growth. Examples of cytokines  
15 representatively include, but are not limited to, interleukins, chemokines, hematopoietic factors (e.g., colony stimulating factors), tumor necrosis factor, and interferons. Representative examples of growth factors include, but are not limited to, platelet-derived growth  
20 factor (PDGF), epidermal growth factor (EGF), fibroblast growth factor (FGF), hepatocyte growth factor (HGF), endothelial cell growth factor (VEGF), cardiotrophin, and the like, which have proliferative activity.

25           The term "hormone" is herein used in its broadest sense in the art, referring to a physiological organic compound which is produced in a particular organ or cell of an animal or plant, and has a physiological effect on an organ apart from the site producing the compound.  
30 Examples of such a hormone include, but are not limited to,

growth hormones, sex hormones, thyroid hormones, and the like. The scope of hormones may overlap partially with that of cytokines.

5                   As used herein, the term "actin-like substance" refers to a substance which interacts directly or indirectly with actin within cells to alter the form or state of actin. Examples of such a substance include, but are not limited to, extracellular matrix proteins (e.g., fibronectin, vitronectin, laminin, etc.), and the like. Such actin-like  
10 substances include substances identified by the following assays. As used herein, interaction with actin is evaluated by visualizing actin with an actin staining reagent (Molecular Probes, Texas Red-X phalloidin) or the like,  
15 followed by microscopic inspection to observe and determine actin aggregation, actin reconstruction or an improvement in cellular outgrowth rate. Such evaluation may be performed quantitatively or qualitatively. Actin-like substances are herein utilized so as to increase  
20 transfection efficiency. An actin-like substance used herein is derived from any organism, including, for example, mammals, such as human, mouse, bovine, and the like.

                  As used herein, the terms "cell adhesion agent",  
25 "cell adhesion molecule", "adhesion agent" and "adhesion molecule" are used interchangeably to refer to a molecule capable of mediating the joining of two or more cells (cell adhesion) or adhesion between a substrate and a cell. In general, cell adhesion molecules are divided into two  
30 groups: molecules involved in cell-cell adhesion

(intercellular adhesion) (cell-cell adhesion molecules) and molecules involved in cell-extracellular matrix adhesion (cell-substrate adhesion) (cell-substrate adhesion molecules). For a method of the present invention, 5 either type of molecule is useful and can be effectively used. Therefore, cell adhesion molecules herein include a substrate protein and a cellular protein (e.g., integrin, etc.) involved in cell-substrate adhesion. A molecule other than a protein can fall within the concept of a cell 10 adhesion molecule as long as it can mediate cell adhesion.

For cell-cell adhesion, cadherin, a number of molecules belonging in an immunoglobulin superfamily (NCAM, L1, ICAM, fasciclin II, III, etc.), selectin, and the like 15 are known, each of which is known to connect cell membranes via a specific molecular reaction.

On the other hand, a major cell adhesion molecule functioning for cell-substrate adhesion is 20 integrin, which recognizes and binds to various proteins contained in extracellular matrices. These cell adhesion molecules are all located on cell membranes and can be regarded as a type of receptor (cell adhesion receptor). Therefore, receptors present on cell membranes can also be 25 used in a method of the present invention. Examples of such a receptor include, but are not limited to,  $\alpha$ -integrin,  $\beta$ -integrin, CD44, syndecan, aggrecan, and the like. Techniques for cell adhesion are well known as described above and as described in, for example, "Saibogaimatorikkusu 30 -Rinsho heno Oyo- [Extracellular matrix -Clinical

Applications-], Medical Review.

It can be determined whether or not a certain molecule is a cell adhesion molecule, by an assay, such as  
5 biochemical quantification (an SDS-PAGE method, a  
labeled-collagen method, etc.), immunological  
quantification (an enzyme antibody method, a fluorescent  
antibody method, an immunohistological study, etc.), a PDR  
method, a hybridization method, or the like, in which a  
10 positive reaction is detected. Examples of such a cell  
adhesion molecule include, but are not limited to, collagen,  
integrin, fibronectin, laminin, vitronectin, fibrinogen,  
immunoglobulin superfamily members (e.g., CD2, CD4, CD8,  
ICM1, ICAM2, VCAM1), selectin, cadherin, and the like. Most  
15 of these cell adhesion molecules transmit an auxiliary  
signal for cell activation into a cell due to intercellular  
interaction as well as cell adhesion. It can be determined  
whether or not such an auxiliary signal can be transmitted  
into a cell, by an assay, such as biochemical quantification  
20 (an SDS-PAGE method, a labeled-collagen method, etc.),  
immunological quantification (an enzyme antibody method,  
a fluorescent antibody method, an immunohistological study,  
etc.), a PDR method, a hybridization method, or the like,  
in which a positive reaction is detected.

25

Examples of cell adhesion molecules include,  
but are not limited to, immunoglobulin superfamily molecules  
(LFA-3, ICAM-1, CD2, CD4, CD8, ICM1, ICAM2, VCAM1, etc.);  
integrin family molecules (LFA-1, Mac-1, gpIIbIIIa, p150,  
30 p95, VLA1, VLA2, VLA3, VLA4, VLA5, VLA6, etc.); selectin

family molecules (L-selectin, E-selectin, P-selectin, etc.), and the like.

As used herein, the term "extracellular matrix protein" refers to a protein constituting an "extracellular matrix". As used herein, the term "extracellular matrix" (ECM) is also called "extracellular substrate" and has the same meaning as commonly used in the art, and refers to a substance existing between somatic cells no matter whether the cells are epithelial cells or non-epithelial cells. Extracellular matrices are involved in supporting tissue as well as in internal environmental structures essential for survival of all somatic cells. Extracellular matrices are generally produced from connective tissue cells. Some extracellular matrices are secreted from cells possessing basal membrane, such as epithelial cells or endothelial cells. Extracellular matrices are roughly divided into fibrous components and matrices filling there between. Fibrous components include collagen fibers and elastin fibers. A basic component of matrices is glycosaminoglycan (acidic mucopolysaccharide), most of which is bound to non-collagenous protein to form a polymer of a proteoglycan (acidic mucopolysaccharide-protein complex). In addition, matrices include glycoproteins, such as laminin of basal membrane, microfibrils around elastin fibers, fibers, fibronectins on cell surfaces, and the like. Specifically differentiated tissue has the same basic structure. For example, in hyaline cartilage, chondroblasts characteristically produce a large amount of cartilage matrices including proteoglycans. In bones, osteoblasts



produce bone matrices which cause calcification. Examples of extracellular matrices for use in the present invention include, but are not limited to, collagen, elastin, proteoglycan, glycosaminoglycan, fibronectin, laminin, 5 elastic fiber, collagen fiber, and the like.

As used herein, the term "receptor" refers to a molecule which is present on cells, within nuclei, or the like, and is capable of binding to an extracellular or 10 intracellular agent where the binding mediates signal transduction. Receptors are typically in the form of proteins. The binding partner of a receptor is usually referred to as a ligand.

15 As used herein, the term "agonist" refers to an agent which binds to the receptor of a certain biologically acting substance (e.g., ligand, etc.), and has the same or similar function as the function of the substance.

20 As used herein, the term "antagonist" refers to a factor which competitively binds to the receptor of a certain biologically acting substance (ligand), and does not produce a physiological action via the receptor. Antagonists include antagonist drugs, blockers, inhibitors, 25 and the like.

(Devices and solid phase supports)

As used herein, the term "device" refers to a part which can constitute the whole or a portion of an 30 apparatus, and comprises a support (preferably, a solid

phase support) and a target substance carried thereon. Examples of such a device include, but are not limited to, chips, arrays, microtiter plates, cell culture plates, Petri dishes, films, beads, and the like.

5

As used herein, the term "support" refers to a material which can fix a substance, such as a biological molecule. Such a support may be made from any fixing material which has a capability of binding to a biological molecule as used herein via covalent or noncovalent bonds, or which may be induced to have such a capability.

Examples of materials used for supports include any material capable of forming a solid surface, such as, without limitation, glass, silica, silicon, ceramics, silicon dioxide, plastics, metals (including alloys), naturally-occurring and synthetic polymers (e.g., polystyrene, cellulose, chitosan, dextran, and nylon), and the like. A support may be formed of layers made of a plurality of materials. For example, a support may be made of an inorganic insulating material, such as glass, quartz glass, alumina, sapphire, forsterite, silicon oxide, silicon carbide, silicon nitride, or the like. A support may be made of an organic material, such as polyethylene, ethylene, polypropylene, polyisobutylene, polyethylene terephthalate, unsaturated polyester, fluorine-containing resin, polyvinyl chloride, polyvinylidene chloride, polyvinyl acetate, polyvinyl alcohol, polyvinyl acetal, acrylic resin, polyacrylonitrile, polystyrene, acetal resin, polycarbonate, polyamide, phenol resin, urea resin,

epoxy resin, melamine resin, styrene-acrylonitrile copolymer, acrylonitrile-butadiene-styrene copolymer, silicone resin, polyphenylene oxide, polysulfone, and the like. Also in the present invention, nitrocellulose film, nylon film, PVDF film, or the like, which are used in blotting, may be used as a material for a support. When a material constituting a support is in the solid phase, such as a support is herein particularly referred to as a "solid phase support". A solid phase support may be herein in the form of a plate, a microwell plate, a chip, a glass slide, a film, beads, a metal (surface), or the like. A support may be uncoated or may be coated.

As used herein, the term "liquid phase" has the same meanings as are commonly understood by those skilled in the art, typically referring to a state in solution.

As used herein, the term "solid phase" has the same meanings as are commonly understood by those skilled in the art, typically referring to a solid state. As used herein, liquid and solid may be collectively referred to as a "fluid".

As used herein, the term "substrate" refers to a material (preferably, solid) which is used to construct a chip or array according to the present invention. Therefore, substrates are included in the concept of plates. Such a substrate may be made from any solid material which has a capability of binding to a biological molecule as used herein via covalent or noncovalent bonds, or which may be

induced to have such a capability.

Examples of materials used for plates and substrates include any material capable of forming a solid surface, such as, without limitation, glass, silica, silicon, ceramics, silicon dioxide, plastics, metals (including alloys), naturally-occurring and synthetic polymers (e.g., polystyrene, cellulose, chitosan, dextran, and nylon), and the like. A support may be formed of layers made of a plurality of materials. For example, a support may be made of an inorganic insulating material, such as glass, quartz glass, alumina, sapphire, forsterite, silicon oxide, silicon carbide, silicon nitride, or the like. A support may be made of an organic material, such as polyethylene, ethylene, polypropylene, polyisobutylene, polyethylene terephthalate, unsaturated polyester, fluorine-containing resin, polyvinyl chloride, polyvinylidene chloride, polyvinyl acetate, polyvinyl alcohol, polyvinyl acetal, acrylic resin, polyacrylonitrile, polystyrene, acetal resin, polycarbonate, polyamide, phenol resin, urea resin, epoxy resin, melamine resin, styrene-acrylonitrile copolymer, acrylonitrile-butadiene-styrene copolymer, silicone resin, polyphenylene oxide, polysulfone, and the like. A material preferable as a substrate varies depending on various parameters such as a measuring device, and can be selected from the above-described various materials as appropriate by those skilled in the art. For transfection arrays, glass slides are preferable. Preferably, such a substrate may be coated with a substance or have a coating.

30

As used herein, the term "coating" in relation to a solid phase support or substrate refers to an act of forming a film of a material on a surface of the solid phase support or substrate, and also refers to a film itself.

5 Coating is performed for various purposes, such as, for example, improvement in the quality of a solid phase support and substrate (e.g., elongation of life span, improvement in resistance to hostile environment, such as resistance to acids, etc.), an improvement in affinity to a substance

10 integrated with a solid phase support or substrate, and the like. Various materials may be used for such coating, including, without limitation, biological substances (e.g., DNA, RNA, protein, lipid, etc.), polymers (e.g., poly-L-lysine, MAS (available from Matsunami Glass,

15 Kishiwada, Japan), and hydrophobic fluorine resin), silane (APS (e.g.,  $\gamma$ -aminopropyl silane, etc.)), metals (e.g., gold, etc.), in addition to the above-described solid phase support and substrate. The selection of such materials is within the technical scope of those skilled in the art and

20 thus can be performed using techniques well known in the art. In one preferred embodiment, such a coating may be advantageously made of poly-L-lysine, silane (e.g., epoxy silane or mercaptosilane, APS ( $\gamma$ -aminopropyl silane), etc.), MAS, hydrophobic fluorine resin, a metal (e.g., gold, etc.).

25 Such a material may be preferably a substance suitable for cells or objects containing cells (e.g., organisms, organs, etc.).

As used herein, the terms "chip" or "microchip"

30 are used interchangeably to refer to a micro-integrated

circuit which has versatile functions and constitutes a portion of a system. Examples of a chip include, but are not limited to, DNA chips, protein chips, and the like.

5                   As used herein, the term "array" refers to a substrate (e.g., a chip, etc.) which has a pattern of a composition containing at least one (e.g., 1000 or more, etc.) target substance (e.g., DNA, proteins, transfection mixtures, etc.), which are arrayed. Among arrays, patterned substrates having a small size (e.g., 10×10 mm, etc.) are particularly referred to as microarrays. The terms "microarray" and "array" are used interchangeably. Therefore, a patterned substrate having a larger size than that which is described above may be referred to as a  
10 microarray. For example, an array comprises a set of desired transfection mixtures fixed to a solid phase surface or a film thereof. An array preferably comprises at least  $10^2$  antibodies of the same or different types, more preferably at least  $10^3$ , even more preferably at least  $10^4$ ,  
15 and still even more preferably at least  $10^5$ . These antibodies are placed on a surface of up to 125×80 mm, more preferably 10×10 mm. An array includes, but is not limited to, a 96-well microtiter plate, a 384-well microtiter plate, a microtiter plate the size of a glass slide, and the like.  
20 A composition to be fixed may contain one or a plurality of types of target substances. Such a number of target substance types may be in the range of from one to the number of spots, including, without limitation, about 10, about 100, about 500, and about 1,000.

30

As described above, any number of target substances (e.g., proteins, such as antibodies) may be provided on a solid phase surface or film, typically including no more than  $10^8$  biological molecules per substrate, in another embodiment no more than  $10^7$  biological molecules, no more than  $10^6$  biological molecules, no more than  $10^5$  biological molecules, no more than  $10^4$  biological molecules, no more than  $10^3$  biological molecules, or no more than  $10^2$  biological molecules. A composition containing more than  $10^8$  biological molecule target substances may be provided on a substrate. In these cases, the size of a substrate is preferably small. Particularly, the size of a spot of a composition containing target substances (e.g., proteins such as antibodies) may be as small as the size of a single biological molecule (e.g., 1 to 2 nm order). In some cases, the minimum area of a substrate may be determined based on the number of biological molecules on a substrate. A composition containing target substances, which are intended to be introduced into cells, are herein typically arrayed on and fixed via covalent bonds or physical interaction to a substrate in the form of spots having a size of 0.01 mm to 10 mm.

"Spots" of biological molecules may be provided on an array. As used herein, the term "spot" refers to a certain set of compositions containing target substances. As used herein, the term "spotting" refers to an act of preparing a spot of a composition containing a certain target substance on a substrate or plate. Spotting may be performed by any method, for example, pipetting or the like,

or alternatively, by using an automatic device. These methods are well known in the art.

As used herein, the term "address" refers to a  
5 unique position on a substrate, which may be distinguished  
from other unique positions. Addresses are appropriately  
associated with spots. Addresses can have any  
distinguishable shape such that substances at each address  
may be distinguished from substances at other addresses  
10 (e.g., optically). A shape defining an address may be, for  
example, without limitation, a circle, an ellipse, a square,  
a rectangle, or an irregular shape. Therefore, the term  
"address" is used to indicate an abstract concept, while  
the term "spot" is used to indicate a specific concept.  
15 Unless it is necessary to distinguish them from each other,  
the terms "address" and "spot" may be herein used  
interchangeably.

The size of each address particularly depends  
20 on the size of the substrate, the number of addresses on  
the substrate, the amount of a composition containing target  
substances and/or available reagents, the size of  
microparticles, and the level of resolution required for  
any method used for the array. The size of each address may  
25 be, for example, in the range of from 1-2 nm to several  
centimeters, though the address may have any size suited  
to an array.

The spatial arrangement and shape which define  
30 an address are designed so that the microarray is suited



to a particular application. Addresses may be densely arranged or sparsely distributed, or subgrouped into a desired pattern appropriate for a particular type of material to be analyzed.

5

Microarrays are widely reviewed in, for example, "Genomu Kino Kenkyu Purotokoru [Genomic Function Research Protocol] (Jikken Igaku Bessatsu [Special Issue of Experimental Medicine], Posuto Genomu Jidai no Jikken Koza 1 [Lecture 1 on Experimentation in Post-genome Era), "Genomu Ikagaku to korekarano Genomu Iryo [Genome Medical Science and Futuristic Genome Therapy (Jikken Igaku Zokan [Special Issue of Experimental Medicine]), and the like.

15

A vast amount of data can be obtained from a microarray. Therefore, data analysis software is important for facilitating correspondence between clones and spots, data analysis, and the like. Such software may be attached to various detection systems (e.g., Ermolaeva O. et al., (1998) Nat. Genet., 20: 19-23). The format of such a database includes, for example, GATC (genetic analysis technology consortium) proposed by Affymetrix.

20

Micromachining for arrays is described in, for example, Campbell, S.A. (1996), "The Science and Engineering of Microelectronic Fabrication", Oxford University Press; Zaut, P.V. (1996), "Micromicroarray Fabrication: a Practical Guide to Semiconductor Processing", Semiconductor Services; Madou, M.J. (1997), "Fundamentals of Microfabrication", CRC15 Press; Rai-Choudhury, P. (1997),

30

"Handbook of Microlithography, Micromachining, & Microfabrication: Microlithography"; and the like, portions related thereto of which are herein incorporated by reference.

5

(Detection)

In cell analysis or determination in the present invention, various detection methods and means can be used as long as they can be used to detect information attributed  
10 to a cell or a substance interacting therewith. Examples of such detection methods and means include, but are not limited to, visual inspection, optical microscopes, confocal microscopes, reading devices using a laser light source, surface plasmon resonance (SPR) imaging, electric  
15 signals, chemical or biochemical markers, which may be used singly or in combination. Examples of such a detecting device include, but are not limited to, fluorescence analyzing devices, spectrophotometers, scintillation counters, CCD, luminometers, and the like. Any means  
20 capable of detecting a biological molecule may be used.

As used herein, the term "marker" refers to a biological agent for indicating a level or frequency of a substance or state of interest. Examples of such a marker  
25 include, but are not limited to, nucleic acids encoding a gene, gene products, metabolic products, receptors, ligands, antibodies, and the like.

Therefore, as used herein, the term "marker" in  
30 relation to a state of a cell refers to an agent (e.g.,

ligands, antibodies, complementary nucleic acids, etc.) interacting with intracellular factors indicating the state of the cell (e.g., nucleic acids encoding a gene, gene products (e.g., mRNA, proteins, posttranscriptionally  
5 modified proteins, etc.), metabolic products, receptors, etc.) in addition to transcription control factors. In the present invention, such a marker may be used to produce a time-lapse profile which is in turn analyzed. Such a marker may preferably interact with a factor of interest. As used  
10 herein, the term "specificity" in relation to a marker refers to a property of the marker which interacts with a molecule of interest to a significantly higher extent than with similar molecules. Such a marker is herein preferably present within cells or may be present outside cells.

15

As used herein, the term "label" refers to a factor which distinguishes a molecule or substance of interest from others (e.g., substances, energy, electromagnetic waves, etc.). Examples of labeling methods  
20 include, but are not limited to, RI (radioisotope) methods, fluorescence methods, biotinylation methods, chemoluminance methods, and the like. When the above-described nucleic acid fragments and complementary oligonucleotides are labeled by fluorescence methods,  
25 fluorescent substances having different fluorescence emission maximum wavelengths are used for labeling. The difference between each fluorescence emission maximum wavelength may be preferably 10 nm or more. Any fluorescent substance which can bind to a base portion of a nucleic acid  
30 may be used, preferably including a cyanine dye (e.g., Cy3

and Cy5 in the Cy Dye™ series, etc.), a rhodamine 6G reagent, N-acetoxy-N2-acetyl amino fluorine (AAF), AAIF (iodine derivative of AAF), and the like. Examples of fluorescent substances having a difference in fluorescence emission maximum wavelength of 10 nm or more include a combination of Cy5 and a rhodamine 6G reagent, a combination of Cy3 and fluorescein, a combination of a rhodamine 6G reagent and fluorescein, and the like. In the present invention, such a label can be used to alter a sample of interest so that the sample can be detected by detecting means. Such alteration is known in the art. Those skilled in the art can perform such alteration using a method appropriate for a label and a sample of interest.

As used herein, the term "interaction" refers to, without limitation, hydrophobic interactions, hydrophilic interactions, hydrogen bonds, Van der Waals forces, ionic interactions, nonionic interactions, electrostatic interactions, and the like.

As used herein, the term "interaction level" in relation to interaction between two substances (e.g., cells, etc.) refers to the extent or frequency of interaction between the two substances. Such an interaction level can be measured by methods well known in the art. For example, the number of cells which are fixed and actually perform an interaction is counted directly or indirectly (e.g., the intensity of reflected light), for example, without limitation, by using an optical microscope, a fluorescence microscope, a phase-contrast microscope, or the like, or

alternatively by staining cells with a marker, an antibody, a fluorescent label or the like specific thereto and measuring the intensity thereof. Such a level can be displayed directly from a marker or indirectly via a label.

5 Based on the measured value of such a level, the number or frequency of genes, which are actually transcribed or expressed in a certain spot, can be calculated.

(Presentation and display)

10 As used herein, the terms "display" and "presentation" are used interchangeably to refer to an act of providing a profile obtained by a method of the present invention, or information derived therefrom, directly or indirectly, or in an information-processed form. Examples

15 of such displayed forms include, but are not limited to, various methods, such as graphs, photographs, tables, animations, and the like. Such techniques are described in, for example, METHODS IN CELL BIOLOGY, VOL. 56, ed. 1998, pp:185-215, A High-Resolution Multimode Digital Microscope

20 System (Sluder & Wolf, Salmon), which discusses application software for automating a microscope and controlling a camera and the design of a hardware device comprising an automated optical microscope, a camera, and a Z-axis focusing device, which can be used herein. Image

25 acquisition by a camera is described in detail in, for example, Inoue and Spring, Video Microscopy, 2d. Edition, 1997, which is herein incorporated by reference.

Real time display can also be performed using

30 techniques well known in the art. For example, after all

images are obtained and stored in a semi-permanent memory, or substantially at the same time as when an image is obtained, the image can be processed with appropriate application software to obtain processed data. For example, data may  
5 be processed by a method for playing back a sequence of images without interruption, a method for displaying images in real time, or a method for displaying images as a "movie" showing irradiating light as changes or continuation on a focal plane.

10

In another embodiment, application software for measurement and presentation typically includes software for setting conditions for applying stimuli or conditions for recording detected signals. With such a measurement and  
15 presentation application, a computer can have a means for applying a stimulus to cells and a means for processing signals detected from cells, and in addition, can control an optically observing means (a SIT camera and an image filing device) and/or a cell culturing means.

20

By inputting conditions for stimulation on a parameter setting screen using a keyboard, a touch panel, a mouse, or the like, it is possible to set the desired complex conditions for stimulation. In addition, various  
25 conditions, such as a temperature for cell culture, pH, and the like, can be set using a keyboard, a mouse, or the like.

A display screen displays a time-lapse profile detected from a cell or information derived therefrom in  
30 real time or after recording. In addition, another recorded

profile or information derived therefrom of a cell can be displayed while being superimposed with a microscopic image of the cell. In addition to recorded information, measurement parameters in recording (stimulation  
5 conditions, recording conditions, display conditions, process conditions, various conditions for cells, temperature, pH, etc.) can be displayed in real time. The present invention may be equipped with a function of issuing an alarm when a temperature or pH departs from the tolerable  
10 range.

On a data analysis screen, it is possible to set conditions for various mathematical analyses, such as Fourier transformation, cluster analysis, FFT analysis,  
15 coherence analysis, correlation analysis, and the like. The present invention may be equipped with a function of temporarily displaying a profile, a function of displaying topography, or the like. The results of these analyses can be displayed while being superimposed with microscopic  
20 images stored in a recording medium.

(Gene introduction)

Any technique may be used herein for introduction of a nucleic acid molecule into cells,  
25 including, for example, transformation, transduction, transfection, and the like. In the present invention, transfection is preferable.

As used herein, the term "transfection" refers  
30 to an act of performing gene introduction or transfection

by culturing cells with genomic DNA, plasmid DNA, viral DNA, viral RNA or the like in a substantially naked form (excluding viral particles), or adding such a genetic material into cell suspension to allow the cells to take  
5 in the genetic material. A gene introduced by transfection is typically expressed within cells in a temporary manner or may be incorporated into cells in a permanent manner.

Such a nucleic acid molecule introduction  
10 technique is well known in the art and commonly used, and is described in, for example, Ausubel F.A. et al., editors, (1988), Current Protocols in Molecular Biology, Wiley, New York, NY; Sambrook J. et al. (1987) Molecular Cloning: A Laboratory Manual, 2nd Ed. and its 3rd Ed., Cold Spring  
15 Harbor Laboratory Press, Cold Spring Harbor, NY; Special issue, Jikken Igaku [Experimental Medicine] "Experimental Methods for Gene introduction & Expression Analysis", Yodo-sha, 1997; and the like. Gene introduction can be confirmed by method as described herein, such as Northern  
20 blotting analysis and Western blotting analysis, or other well-known, common or routine techniques.

When a gene is mentioned herein, the term "vector" or "recombinant vector" refers to a vector  
25 transferring a polynucleotide sequence of interest to a target cell. Such a vector is capable of self-replication or incorporation into a chromosome in a host cell (e.g., a prokaryotic cell, yeast, an animal cell, a plant cell, an insect cell, an individual animal, and an individual plant,  
30 etc.), and contains a promoter at a site suitable for



transcription of a polynucleotide of the present invention. A vector suitable for performing cloning is referred to as a "cloning vector". Such a cloning vector ordinarily contains a multiple cloning site containing a plurality of  
5 restriction sites. Restriction enzyme sites and multiple cloning sites as described above are well known in the art and can be used as appropriate by those skilled in the art depending on the purpose in accordance with publications described herein (e.g., Sambrook et al., *supra*).

10

As used herein, the term "expression vector" refers to a nucleic acid sequence comprising a structural gene and a promoter for regulating expression thereof, and in addition, various regulatory elements in a state that  
15 allows them to operate within host cells. The regulatory element may include, preferably, terminators, selectable markers such as drug-resistance genes, and enhancers.

Examples of "recombinant vectors" for  
20 prokaryotic cells include, but are not limited to, pCDNA3(+), pBluescript-SK(+/-), pGEM-T, pEF-BOS, pEGFP, pHAT, pUC18, pFT-DEST<sup>TM</sup>42GATEWAY (Invitrogen), and the like.

Examples of "recombinant vectors" for animal  
25 cells include, but are not limited to, pCDNAI/Amp, pCDNAI, pCDM8 (all commercially available from Funakoshi), pAGE107 [Japanese Laid-Open Publication No. 3-229 (Invitrogen), pAGE103 [J. Biochem., 101, 1307(1987)], pAMo, pAMoA [J. Biol. Chem., 268, 22782-22787(1993)], a retrovirus expression  
30 vector based on a murine stem cell virus (MSCV), pEF-BOS,

pEGFP, and the like.

Examples of recombinant vectors for plant cells include, but are not limited to, pPCVICEn4HPT, pCGN1548,  
5 pCGN1549, pBI221, pBI121, and the like.

Any of the above-described methods for introducing DNA into cells can be used as a vector introduction method, including, for example, transfection,  
10 transduction, transformation, and the like (e.g., a calcium phosphate method, a liposome method, a DEAE dextran method, an electroporation method, a particle gun (gene gun) method, and the like), a lipofection method, a spheroplast method (Proc. Natl. Acad. Sci. USA, 84, 1929(1978)), a lithium  
15 acetate method (J. Bacteriol., 153, 163(1983); and Proc. Natl. Acad. Sci. USA, 75, 1929(1978)), and the like.

As used herein, the term "operably linked" indicates that a desired sequence is located such that  
20 expression (operation) thereof is under control of a transcription and translation regulatory sequence (e.g., a promoter, an enhancer, and the like) or a translation regulatory sequence. In order for a promoter to be operably linked to a gene, typically, the promoter is located  
25 immediately upstream of the gene. A promoter is not necessarily adjacent to a structural gene.

As used herein, the term "gene introduction reagent" refers to a reagent which is used in a gene  
30 introduction method so as to enhance introduction efficiency.

Examples of such a gene introduction reagent include, but are not limited to, cationic polymers, cationic lipids, polyamine-based reagents, polyimine-based reagents, calcium phosphate, and the like. Specific examples of a reagent used in transfection include reagents available from various sources, such as, without limitation, Effectene Transfection Reagent (cat. no. 301425, Qiagen, CA), TransFast™ Transfection Reagent (E2431, Promega, WI), Tfx™-20 Reagent (E2391, Promega, WI), SuperFect Transfection Reagent (301305, Qiagen, CA), PolyFect Transfection Reagent (301105, Qiagen, CA), LipofectAMINE 2000 Reagent (11668-019, Invitrogen corporation, CA), JetPEI (×4) conc. (101-30, Polyplus-transfection, France) and ExGen 500 (R0511, Fermentas Inc., MD), and the like.

Gene expression (e.g., mRNA expression, polypeptide expression) may be "detected" or "quantified" by an appropriate method, including mRNA measurement and immunological measurement methods. Examples of molecular biological measurement methods include Northern blotting methods, dot blotting methods, PCR methods, and the like. Examples of immunological measurement methods include ELISA methods, RIA methods, fluorescent antibody methods, Western blotting methods, immunohistological staining methods, and the like, where a microtiter plate may be used. Examples of quantification methods include ELISA methods, RIA methods, and the like. A gene analysis method using an array (e.g., a DNA array, a protein array, etc.) may be used. The DNA array is widely reviewed in Saibo-Kogaku [Cell Engineering], special issue, "DNA Microarray and Up-to-date PCR Method",

edited by Shujun-sha. The protein array is described in detail in Nat Genet. 2002 Dec; 32 Suppl:526-32. Examples of methods for analyzing gene expression include, but are not limited to, RT-PCR methods, RACE methods, SSCP methods, immunoprecipitation methods, two-hybrid systems, *in vitro* translation methods, and the like, in addition to the above-described techniques. Other analysis methods are described in, for example, "Genome Analysis Experimental Method, Yusuke Nakamura's Lab-Manual, edited by Yusuke Nakamura, Yodosha (2002), and the like. All of the above-described publications are herein incorporated by reference.

As used herein, the term "expression level" refers to the amount of a polypeptide or mRNA expressed in a subject cell. The term "expression level" includes the level of protein expression of a polypeptide evaluated by any appropriate method using an antibody, including immunological measurement methods (e.g., an ELISA method, an RIA method, a fluorescent antibody method, a Western blotting method, an immunohistological staining method, and the like, or the mRNA level of expression of a polypeptide evaluated by any appropriate method, including molecular biological measurement methods (e.g., a Northern blotting method, a dot blotting method, a PCR method, and the like). The term "change in expression level" indicates that an increase or decrease in the protein or mRNA level of expression of a polypeptide evaluated by an appropriate method including the above-described immunological measurement methods or molecular biological measurement

methods.

(Screening)

As used herein, the term "screening" refers to  
5 selection of a target, such as an organism, a substance,  
or the like, a given specific property of interest from a  
population containing a number of elements using a specific  
operation/evaluation method. For screening, an agent (e.g.,  
an antibody), a polypeptide or a nucleic acid molecule of  
10 the present invention can be used.

As used herein, screening by utilizing an  
immunological reaction is also referred to as  
"immunophenotyping". In this case, an antibody or a single  
15 chain antibody may be used for immunophenotyping a cell line  
and a biological sample. A transcription or translation  
product of a gene may be useful as a cell specific marker,  
or more particularly, a cell marker which is distinctively  
expressed in various stages in differentiation and/or  
20 maturation of a specific cell type. A monoclonal antibody  
directed to a specific epitope, or a combination of epitopes  
allows for screening of a cell population expressing a marker.  
Various techniques employ monoclonal antibodies to screen  
for a cell population expressing a marker. Examples of such  
25 techniques include, but are not limited to, magnetic  
separation using magnetic beads coated with antibodies,  
"panning" using antibodies attached to a solid matrix (i.e.,  
a plate), flow cytometry, and the like (e.g., US Patent  
No. 5,985,660; and Morrison et al., Cell, 96:737-49 (1999)).

30

These techniques may be used to screen cell populations containing undifferentiated cells, which can grow and/or differentiate as seen in human umbilical cord blood or which are treated and modified into an undifferentiated state (e.g., embryonic stem cells, tissue stem cells, etc.).

#### (Diagnosis)

As used herein, the term "diagnosis" refers to an act of identifying various parameters associated with a disease, a disorder, a condition, or the like of a subject and determining a current state of the disease, the disorder, the condition, or the like. A method, device, or system of the present invention can be used to analyze a sugar chain structure, a drug resistance level, or the like. Such information can be used to select parameters, such as a disease, a disorder, a condition, and a prescription or method for treatment or preventative method for a subject.

A diagnosis method of the present invention can use, in principle, a sample which is derived from the body of a subject. Therefore, it is possible for some one which is not a medical practitioner, such as a medical doctor, to deal with such a sample. The present invention is industrially useful.

#### (Therapy)

As used herein, the term "therapy" refers to an act of preventing progression of a disease or a disorder, preferably maintaining the current state of a disease or

a disorder, more preferably alleviating a disease or a disorder, and more preferably extinguishing a disease or a disorder.

5           As used herein, the term "subject" refers to an organism which is subjected to the treatment of the present invention. A subject is also referred to as a "patient". A patient or subject may preferably be a human.

10           As used herein, the term "cause" or "pathogen" in relation to a disease, a disorder or a condition of a subject refers to an agent associated with the disease, the disorder or the condition (also collectively referred to as a "lesion", or "disease damage" in plants), including,  
15 without limitation, a causative or pathogenic substance (pathogenic agent), a disease agent, a disease cell, a pathogenic virus, and the like.

          A disease targeted by the present invention may  
20 be any disease associated with a pathogenic gene. Examples of such a disease include, but are not limited to, cancer, infectious diseases due to viruses or bacteria, allergy, hypertension, hyperlipemia, diabetes, cardiac diseases, cerebral infarction, dementia, obesity, arteriosclerosis,  
25 infertility, mental and nervous diseases, cataract, progeria, hypersensitivity to ultraviolet radiation, and the like.

          A disorder targeted by the present invention may  
30 be any disorder associated with a pathogenic gene.

Examples of such a disease, disorder or condition include, but are not limited to, circulatory diseases (anemia (e.g., aplastic anemia (particularly, severe aplastic anemia), renal anemia, cancerous anemia, secondary anemia, refractory anemia, etc.), cancer or tumors (e.g., leukemia, multiple myeloma), etc.); neurological diseases (dementia, cerebral stroke and sequela thereof, cerebral tumor, spinal injury, etc.); immunological diseases (T-cell deficiency syndrome, leukemia, etc.); motor organ and the skeletal system diseases (fracture, osteoporosis, luxation of joints, subluxation, sprain, ligament injury, osteoarthritis, osteosarcoma, Ewing's sarcoma, osteogenesis imperfecta, osteochondrodysplasia, etc.); dermatologic diseases (atrachia, melanoma, cutis malignant lymphoma, hemangiosarcoma, histiocytosis, hydroa, pustulosis, dermatitis, eczema, etc.); endocrinologic diseases (hypothalamus/hypophysis diseases, thyroid gland diseases, accessory thyroid gland (parathyroid) diseases, adrenal cortex/medulla diseases, saccharometabolism abnormality, lipid metabolism abnormality, protein metabolism abnormality, nucleic acid metabolism abnormality, inherent metabolic disorders (phenylketonuria, galactosemia, homocystinuria, maple syrup urine disease), analbuminemia, lack of ascorbic acid synthetic ability, hyperbilirubinemia, hyperbilirubinuria, kallikrein deficiency, mast cell deficiency, diabetes insipidus, vasopressin secretion abnormality, dwarfism, Wolman's disease (acid lipase deficiency, mucopolysaccharidosis VI, etc.); respiratory diseases (pulmonary diseases (e.g.,



pneumonia, lung cancer, etc.), bronchial diseases, lung cancer, bronchial cancer, etc.); alimentary diseases (esophageal diseases (e.g., esophagial cancer, etc.), stomach/duodenum diseases (e.g., stomach cancer, duodenum cancer, etc.), small intestine diseases/large intestine diseases (e.g., polyps of the colon, colon cancer, rectal cancer, etc.), bile duct diseases, liver diseases (e.g., liver cirrhosis, hepatitis (A, B, C, D, E, etc.), fulminant hepatitis, chronic hepatitis, primary liver cancer, alcoholic liver disorders, drug induced liver disorders, etc.), pancreatic diseases (acute pancreatitis, chronic pancreatitis, pancreas cancer, cystic pancreas diseases, etc.), peritoneum/abdominal wall/diaphragm diseases (hernia, etc.), Hirschsprung's disease, etc.); urinary diseases (kidney diseases (e.g., renal failure, primary glomerulus diseases, renovascular disorders, tubular function abnormality, interstitial kidney diseases, kidney disorders due to systemic diseases, kidney cancer, etc.), bladder diseases (e.g., cystitis, bladder cancer, etc.); genital diseases (male genital organ diseases (e.g., male sterility, prostatomegaly, prostate cancer, testicular cancer, etc.), female genital organ diseases (e.g., female sterility, ovary function disorders, hystero myoma, adenomyosis uteri, uterine cancer, endometriosis, ovarian cancer, villosity diseases, etc.), etc); circulatory diseases (heart failure, angina pectoris, myocardial infarct, arrhythmia, valvulitis, cardiac muscle/pericardium diseases, congenital heart diseases (e.g., atrial septal defect, arterial canal patency, tetralogy of Fallot, etc.), artery diseases (e.g.,

arteriosclerosis, aneurysm), vein diseases (e.g., phlebeurysm, etc.), lymphoduct diseases (e.g., lymphedema, etc.), etc.); and the like.

5           As used herein, the term "cancer" refers to a malignant tumor which has a high level of atypism, grows faster than normal cells, tends to disruptively invade surrounding tissue or metastasize to new body sites or a condition characterized by the presence of such a malignant  
10 tumor. In the present invention, cancer includes, without limitation, solid cancer and hematological cancer.

          As used herein, the term "solid cancer" refers to a cancer having a solid shape in contrast to hematological  
15 cancer, such as leukemia and the like. Examples of such a solid cancer include, but are not limited to, breast cancer, liver cancer, stomach cancer, lung cancer, head and neck cancer, uterocervical cancer, prostate cancer, retinoblastoma, malignant lymphoma, esophagus cancer,  
20 brain tumor, osteoncus, and the like.

          As used herein, the term "cancer therapy" encompasses administration of an anticancer agent (e.g., a chemotherapeutic agent, radiation therapy, etc.) or  
25 surgical therapy, such as surgical excision and the like.

          Chemotherapeutic agents used herein are well known in the art and are described in, for example, Shigeru Tsukagoshi et al. editors, "Kogan zai Manuaru [Manual of  
30 Anticancer agents]", 2nd ed., Chugai Igaku sha;

Pharmacology; and Lippincott Williams & Wilkins, Inc. Examples of such chemotherapeutic agents are described below: 1) alkylating agents which alkylate cell components, such as DNA, protein, and the like, to produce cytotoxicity (e.g., cyclophosphamide, busulfan, thiotepa, dacarbazine, etc.); 2) antimetabolites which mainly inhibit synthesis of nucleic acids (e.g., antifolates (methotrexate, etc.), antipurines (6-mercaptopurine, etc.), antipyrimidines (fluorouracil (5-FU), etc.); 3) DNA topoisomerase inhibitors (e.g., camptothecin and etoposide, each of which inhibits topoisomerases I and II)); 4) tubulin agents which inhibit formation of microtubules and suppress cell division (vinblastine, vincristine, etc.); 5) platinum compounds which bind to DNA and proteins to exhibit cytotoxicity (cisplatin, carboplatin, etc.); 6) anticancer antibiotics which bind to DNA to inhibit synthesis of DNA and RNA (adriamycin, dactinomycin, mitomycin C, bleomycin, etc.); 7) hormone agents which are applicable to hormone-dependent cancer, such as breast cancer, uterus cancer, prostate cancer, and the like (e.g., tamoxifen, leuporelin (LH-RH), etc.); 8) biological formulations (asparaginase effective for asparagine requiring blood malignant tumor, interferon exhibiting direct antitumor action and indirect action by immunopotentialization, etc.); 9) immunostimulants which exhibit capability of immune response, indirectly leading to antitumor activity (e.g., rentinan which is a polysaccharide derived from shiitake mushroom, bestatin which is a peptide derived from a microorganism, etc.).

An "anticancer agent" used herein selectively

suppresses the growth of cancerous (tumor) cells, and includes both pharmaceutical agents and radiation therapy. Such an anticancer agent is well known in the art and described in, for example, Shigeru Tsukagoshi et al. editors, 5 "Kogan zai Manuaru [Manual of Anticancer agents]", 2nd ed., Chugai Igaku sha; Pharmacology; and Lippincott Williams & Wilkins, Inc.

As used herein, the term "radiation therapy" 10 refers to a therapy for diseases using ionizing radiation or radioactive substances. Representative examples of radiation therapy include, but are not limited to, X-ray therapy,  $\gamma$ -ray therapy, electron beam therapy, proton beam therapy, heavy particle beam therapy, neutron capture 15 therapy, and the like. For example, heavy particle beam therapy is preferable. However, heavy particle beam therapy requires a large-size device and is not generally used. The above-described radiation therapies are well known in the art and are described in, for example, Sho Kei 20 Zen, "Hoshasen kensa to Chiryo no Kiso: Hoshasen Chiryo to Shugakuteki Chiryo [Basics of Radiation Examination and Therapies: Radiation Therapy and Incentive Therapy]", (Shiga Medical School, Radiation): Total digestive system care, Vol. 6, No. 6, Pages 79-89, 6-7 (2002.02). For drug 25 resistance to be identified in the present invention, chemotherapies are typically considered. However, resistance to radiation therapy is also associated with time-lapse profiles. Therefore, radiation therapy is herein encompassed by the concept of pharmaceutical agents.

30

As used herein, the term "pharmaceutically acceptable carrier" refers to a material for use in production of a medicament, an animal drug or an agricultural chemical, which does not have an adverse effect on an effective component. Examples of such a pharmaceutically acceptable carrier include, but are not limited to, antioxidants, preservatives, colorants, flavoring agents, diluents, emulsifiers, suspending agents, solvents, fillers, bulking agents, buffers, delivery vehicles, excipients, agricultural or pharmaceutical adjuvants, and the like.

The type and amount of a pharmaceutical agent used in a treatment method of the present invention can be easily determined by those skilled in the art based on information obtained by a method of the present invention (e.g., information about the level of drug resistance, etc.) and with reference to the purpose of use, a target disease (type, severity, and the like), the patient's age, weight, sex, and case history, the form or type of the cell, and the like. The frequency of the treatment method of the present invention applied to a subject (or patient) is also determined by those skilled in the art with respect to the purpose of use, target disease (type, severity, and the like), the patient's age, weight, sex, and case history, the progression of the therapy, and the like. Examples of the frequency include once per day to several months (e.g., once per week to once per month). Preferably, administration is performed once per week to once per month with reference to the progression of the therapy.

As used herein, the term "instructions" refers to a description of a tailor made therapy of the present invention for a person who performs administration, such as a medical doctor, a patient, or the like. Instructions state when to administer a medicament of the present invention, such as immediately after or before radiation therapy (e.g., within 24 hours, etc.). The instructions are prepared in accordance with a format defined by an authority of a country in which the present invention is practiced (e.g., Health, Labor and Welfare Ministry in Japan, Food and Drug Administration (FDA) in the U.S., and the like), explicitly describing that the instructions are approved by the authority. The instructions are so-called package insert and are typically provided in paper media. The instructions are not so limited and may be provided in the form of electronic media (e.g., web sites, electronic mails, and the like provided on the internet).

In a therapy of the present invention, two or more pharmaceutical agents may be used as required. When two or more pharmaceutical agents are used, these agents may have similar properties or may be derived from similar origins, or alternatively, may have different properties or may be derived from different origins. A method of the present invention can be used to obtain information about the drug resistance level of a method of administering two or more pharmaceutical agents.

Also, in the present invention, gene therapy can

be performed based on the resultant information about drug resistance. As used herein, the term "gene therapy" refers to a therapy in which a nucleic acid, which has been expressed or can be expressed, is administered into a subject. In such  
5 an embodiment of the present invention, a protein encoded by a nucleic acid is produced to mediate a therapeutic effect.

In the present invention, it will be understood  
10 by those skilled in the art that if the result of analysis of a certain specific time-lapse profile is correlated with a state of a cell in a similar organism (e.g., mouse with respect to human, etc.), the result of an analysis of a corresponding time-lapse profile can be correlated with a  
15 state of a cell. This feature is supported by, for example, Dobutsu Baiyo Saibo Manuaru [Animal Culture Cell Manual], Seno, ed., Kyoritsu Shuppan, 1993, which is herein incorporated by reference.

20 The present invention may be applied to gene therapies. As used herein, the term "gene therapy" refers to a therapy in which a nucleic acid, which has been expressed or can be expressed, is administered into a subject. In such an embodiment of the present invention, a protein encoded  
25 by a nucleic acid is produced to mediate a therapeutic effect.

Any methods for gene therapy available in the art may be used in accordance with the present invention.  
30 Illustrative methods will be described below.

Methods for gene therapy are generally reviewed in, for example, Goldspiel et al., Clinical Pharmacy 12: 488-505(1993); Wu and Wu, Biotherapy 3: 87-95(1991);  
5 Tolstoshev, Ann. Rev. Pharmacol. Toxicol., 32: 573-596(1993); Mulligan, Science 260: 926-932(1993); Morgan and Anderson, Ann. Rev. Biochem., 62: 191-217(1993); and May, TIBTECH 11(5): 155-215(1993). Commonly known recombinant DNA techniques used in gene therapy are  
10 described in, for example, Ausubel et al. (ed.), Current Protocols in Molecular Biology, John Wiley & Sons, NY(1993); and Kriegler, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY (1990).

15 (Basic techniques)

Techniques used herein are within the technical scope of the present invention unless otherwise specified. These techniques are commonly used in the fields of fluidics, micromachining, organic chemistry, biochemistry, genetic  
20 engineering, molecular biology, microbiology, genetics, and their relevant fields. The techniques are well described in documents described below and the documents mentioned herein elsewhere.

25 Microfabrication is described in, for example, Campbell, S.A. (1996), "The Science and Engineering of Microelectronic Fabrication", Oxford University Press; Zaut, P.V. (1996), "Micromicroarray Fabrication: a Practical Guide to Semiconductor Processing",  
30 Semiconductor Services; Madou, M.J. (1997), "Fundamentals



of Microfabrication", CRC15 Press; Rai-Choudhury, P. (1997),  
 "Handbook of Microlithography, Micromachining, &  
 Microfabrication: Microlithography". Relevant portions  
 (or possibly the entirety) of each of these publications  
 5 are herein incorporated by reference.

Molecular biology techniques, biochemistry  
 techniques, and microbiology techniques used herein are well  
 known and commonly used in the art, and are described in,  
 10 for example, Sambrook J. et al. (1989), "Molecular Cloning:  
 A Laboratory Manual", Cold Spring Harbor and its 3rd Ed.  
 (2001); Ausubel, F.M. (1987), "Current Protocols in  
 Molecular Biology", Greene Pub. Associates and  
 Wiley-Interscience; Ausubel, F.M. (1989), "Short Protocols  
 15 in Molecular Biology: A Compendium of Methods from Current  
 Protocols in Molecular Biology", Greene Pub. Associates and  
 Wiley-Interscience; Innis, M.A. (1990), "PCR Protocols:  
 A Guide to Methods and Applications", Academic Press;  
 Ausubel, F.M. (1992), "Short Protocols in Molecular Biology:  
 20 A Compendium of Methods from Current Protocols in Molecular  
 Biology", Greene Pub. Associates; Ausubel, F.M. (1995),  
 "Short Protocols in Molecular Biology: A Compendium of  
 Methods from Current Protocols in Molecular Biology", Greene  
 Pub. Associates; Innis, M.A. et al. (1995), "PCR  
 25 Strategies", Academic Press; Ausubel, F.M. (1999), "Short  
 Protocols in Molecular Biology: A Compendium of Methods  
 from Current Protocols in Molecular Biology", Wiley, and  
 annual updates; Sninsky, J.J. et al. (1999), "PCR  
 Applications: , Protocols for Functional Genomics",  
 30 Academic Press; Special issue, Jikken Igaku [Experimental

Medicine] "Idenshi Donyu & Hatsugenkaiseki Jikkenho  
[Experimental Method for Gene introduction & Expression  
Analysis]", Yodo-sha, 1997; and the like. Relevant  
portions (or possibly the entirety) of each of these  
5 publications are herein incorporated by reference.

DNA synthesis techniques and nucleic acid  
chemistry for producing artificially synthesized genes are  
described in, for example, Gait, M.J. (1985),  
10 "Oligonucleotide Synthesis: A Practical Approach", IRL  
Press; Gait, M.J. (1990), "Oligonucleotide Synthesis: A  
Practical Approach", IRL Press; Eckstein, F. (1991),  
"Oligonucleotides and Analogues: A Practical Approach", IRL  
Press; Adams, R.L. et al. (1992), "The Biochemistry of the  
15 Nucleic Acids", Chapman & Hall; Shabarova, Z. et al. (1994),  
"Advanced Organic Chemistry of Nucleic Acids", Weinheim;  
Blackburn, G.M. et al. (1996), "Nucleic Acids in Chemistry  
and Biology", Oxford University Press; Hermanson, G.T.  
(1996), "Bioconjugate Techniques", Academic Press; and the  
20 like. Relevant portions (or possibly the entirety) of each  
of these publications are herein incorporated by reference.

(Analysis of co-regulation of genes)

Mathematical processes used herein can be  
25 performed by using well-known techniques described in, for  
example, Kazuyuki Shimizu, "Seimei Sisutemu Kaiseki  
notameno Sugaku [Mathematics for Analyzing Biological  
Systems]", Corona sha, 1999; and the like. Among these  
techniques, representative analysis techniques will be  
30 described below.

In one embodiment, such a mathematical process may be regression analysis. Examples of regression analysis include, but are not limited to, linear regression (e.g., simple regression analysis, multiple regression analysis, robust estimation, etc.), nonlinear estimation, and the like.

In simple regression analysis,  $n$  sets of data  $(x_1, y_1)$  to  $(x_n, y_n)$  are fitted to  $y_i = ax_i + b + e_i$  ( $i=1, 2, \dots, n$ ) where  $a$  and  $b$  are model parameters, and  $e_i$  represents a deviation or an error from the straight line. The parameters  $a$  and  $b$  are typically determined so that the mean of a sum of squares of the distance between a data point and the straight line is minimal. In this case, the rms of the distance is partially differentiated to produce simultaneous linear equations. These equations are solved for  $a$  and  $b$  which minimize the square errors. Such values are called least square estimates.

20

Next, a regression line is calculated based on the value obtained by subtracting the mean of all data values from each data value. A regression line represented by:

25

$$A \sum_i X_i + B = \sum Y_i$$

is assumed. Further, it is assumed that  $B=0$ . The mean  $(x_{ave}, y_{ave})$  of  $(x_i, y_i)$  ( $i=1, 2, \dots, n$ ) is calculated, and the variance of  $x$  ( $s_{xx}$ ) and the covariance of  $x$  and  $y$  ( $s_{xy}$ ) are calculated.

The above-described regression line can be represented by:

$$Y - Y_{ave} = (S_{xy}/S_{xx})(x - x_{ave}).$$

The correlation coefficient  $r_{xy}$  is represented by:

5

$$r_{xy} = s_{xy}/\sqrt{(s_{xy}s_{yy})}.$$

In this case, the relationship  $\Sigma e_i^2/n = s_{yy}(1 - r_{xy}^2)$  is satisfied. Therefore, as  $|r_{xy}|$  approaches 1, the  
 10 error is decreased, which means that data can be satisfactorily represented by the regression line.

In another embodiment, multiple regression analysis is used. In this technique,  $y$  is not a single  
 15 independent variable, and is considered to be a function of two or more variables, e.g., is represented by:

$$Y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n.$$

20 This equation is called a multiple regression equation.  $a_0$  and the like are called (partial) regression coefficients. In multiple regression analysis, a least square method is used and normal equations are solved to obtain least square estimates. Evaluation can be performed  
 25 as with single regression analysis.

In another embodiment, robust estimation is used. The least square method is based on the premise that measurement values are not biased and measurement errors  
 30 have a normal distribution, and models have no approximation

error. In actual situations, however, there may be errors in measurement. In robust estimation, unreliable data is detected and separated as outliers from the great majority of data which are reliable, or is subjected to a statistical process. Such a robust estimation may be utilized herein.

Nonlinear estimation may also be used herein. With nonlinear estimation, it is possible to represent a nonlinear model as vector equations which are in turn solved.

10

Other mathematical processes used herein include principal component analysis, which utilizes two-dimensional data principal component analysis, multi-dimensional data principal component analysis, singular value decomposition, and generalized inverse matrix. Alternatively, canonical correlation analysis, factor analysis, discrimination analysis, cluster analysis, and the like may be used herein.

(Gene set classification by cluster analysis)

For a number of applications, it may be desirable to obtain a set of reference transcription control sequences which are cooperatively controlled under a wide range of conditions. An embodiment of identifying such a set of reference transcription control sequences is, for example, a clustering algorithm, which is reviewed in, for example, Fukunaga, 1990, "Statistical Pattern Recognition", 2nd ed., Academic Press, San Diego; Anderberg, 1973, "Cluster Analysis for Applications", Academic Press: New York; Everitt, 1974, "Cluster Analysis", London: Heinemann

Educ. Books; Hartigan, 1975, "Clustering Algorithms", New York: Wiley; and Sneath and Sokal, 1973, "Numerical Taxonomy", Freeman.

5                   A set of transcription control sequences can also be defined based on a transcription control mechanism. Transcription control sequences having a transcription factor binding site for the same or similar sequences in a regulatory region are likely to be cooperatively regulated.

10 In a certain embodiment, the regulatory regions of transcription control sequences of interest are compared with one another using multiple alignment analysis, so that a possible common transcription factor binding site can be determined (Stormo and Hartzell, 1989, "Identifying protein

15 binding sites from unaligned DNA fragments", Proc. Natl. Acad. Sci., 86: 1183-1187; Hertz and Stormo, 1995, "Identification of consensus patterns in unaligned DNA and protein sequences: a large-deviation statistical basis for penalizing gaps", Proc. of 3rd Intl. Conf. on Bioinformatics

20 and Genome Research, Lim and Cantor, ed., World Scientific Publishing Co., Ltd. Singapore, pp.201-216).

                  It may be desirable to obtain a set of basic transcription control sequences which are cooperatively

25 regulated under various conditions. With such a set, a method of the present invention can satisfactorily and efficiently carry out determination based on profiles. A preferable embodiment for identifying such a set of basic transcription control sequences includes a clustering

30 algorithm.

In an embodiment using cluster analysis, the transcription levels of a number of transcription control sequences can be monitored while applying various stimuli to biological samples. A table of data containing measurements of the transcription levels of transcription control sequences is used in cluster analysis. In order to obtain a set of basic transcription control sequences containing transcription control sequences which simultaneously vary under various conditions, typically at least two, preferably at least 3, more preferably at least 10, even more preferably more than 50, and most preferably more than 100 stimuli or conditions are used. Cluster analysis is performed for a table of data having  $m \times k$  dimensions where  $m$  is the total number of conditions or stimuli and  $k$  is the number of transcription control sequences to be measured.

A number of clustering algorithms are useful for clustering analysis. In clustering algorithms, differences or distances between samples are used to form clusters. In a certain embodiment, a distance used is a Euclidean distance

$$I(x,y)=\left\{\sum_i (X_i - Y_i)^2\right\}^{1/2} \quad \text{in space:}$$

(1)

where  $(x, y)$  represents a distance between gene  $X$  and gene  $Y$  (or any other cellular components  $X$  and  $Y$  (e.g., transcription control sequences));  $X_i$  and  $Y_i$  represent gene expression in response to  $i$  stimuli. Euclidean distances

may be squared and then multiplied with weighting which are increased with an increase in the distance. Alternatively, a distance reference may be, for example, a distance between transcription control sequences X and Y, or a Manhattan distance represented by

$$I(x,y) = \sum_i |X_i - Y_i| \quad (2)$$

where  $X_i$  and  $Y_i$  represent responses of transcription control sequences or gene expression when  $i$  stimuli are applied. Several other definitions of distance include Chebyshev distance, power distance, and mismatch rate. When dimensional data can be categorized without modification, a mismatch rate defined as  $I(x, y) = (\text{the number of } X_i \neq Y_i) / i$  may be used in a method of the present invention. Such a method is particularly useful in terms of cellular responses. Another useful definition of distance is  $I=1-r$  where  $r$  is a correlation coefficient of response vectors  $X$  and  $Y$ , e.g., a normalized inner product  $X \cdot Y / |X| |Y|$ . Specifically, an inner product  $X \cdot Y$

$$X \cdot Y = \sum_i X_i \times Y_i \quad (3)$$

Also,  $|X| = (X \cdot X)^{1/2}$  and  $|Y| = (Y \cdot Y)^{1/2}$ .

Most preferably, a distance reference is suited to a biological problem in order to identify cellular components (e.g., transcription control sequences, etc.) which are simultaneously changed and/or simultaneously regulated. For example, in a particularly preferred



embodiment, correlation coefficient of genes X and Y in experiment i, r

$$r = \frac{\sum_i \frac{X_i Y_i}{\sigma_i^{(X)} \sigma_i^{(Y)}}}{\left[ \sum_i \left( \frac{X_i}{\sigma_i^{(X)}} \right)^2 \left( \frac{Y_i}{\sigma_i^{(Y)}} \right)^2 \right]^{1/2}}$$

I=1-r having a weighted inner product such a preferred embodiment, r

(4)

where  $\sigma_i^{(X)}$  and  $\sigma_i^{(Y)}$  represent standard errors in measurement of genes X and Y in experiment i.

10

The above-described normalized and weighted inner products (correlation coefficients) are constrained between values +1 (two response vectors are completely correlated, i.e., the two vectors are essentially the same) and -1 (two response vectors are not correlated or do not have the same orientation (i.e., opposing orientations)). These correlation coefficients are particularly preferable in an embodiment of the present invention which tries to detect a set or cluster of cellular components (e.g., transcription control sequences, etc.) having the same sign or response.

In another embodiment, it is preferable to identify a set or cluster of cellular components (e.g., transcription control sequences, etc.) which simultaneously regulate the same biological response or pathway or are involved in such regulation, or have similar or non-correlated responses. In such a embodiment, it is preferable to use the absolute value of either the above-described normalized or weighted inner product, i.e.,

30

$|r|$  as a correlation coefficient.

In still another embodiment, the relationship between cellular components (e.g., transcription control sequences, etc.), which are simultaneously regulated and/or simultaneously changed, are more complicated, e.g., a number of biological pathways (e.g., signal transduction pathways, etc.) are involved with the same cellular component (e.g., a transcription control sequence, etc.) so that different results may be obtained. In such an embodiment, it is preferable to use a correlation coefficient  $r=r^{(\text{change})}$  which can identify cellular components (other transcription control sequences as controls which are not involved in change) when simultaneous represented the above-de

$$r = \frac{\sum_i \left| \frac{x_i}{\sigma_i^{(x)}} \right| \left| \frac{y_i}{\sigma_i^{(y)}} \right|}{\left[ \sum_i \left( \frac{x_i}{\sigma_i^{(x)}} \right)^2 \left( \frac{y_i}{\sigma_i^{(y)}} \right)^2 \right]^{1/2}}$$

(5).

Various cluster linkage methods are useful in a method of the present invention.

Examples of such a technique include a simple linkage method, a nearest neighbor method, and the like. In these techniques, a distance between the two closest samples is measured. Alternatively, in a complete linkage method, which may be herein used, a maximum distance between two samples in different clusters is measured. This technique is particularly useful when genes or other

cellular components naturally form separate "clumps".

Alternatively, the mean of non-weighted pairs is used to define the mean distance of all sample pairs in two different clusters. This technique is also useful in clustering genes or other cellular components which naturally form separate "clumps". Finally, a weighted pair mean technique is also available. This technique is the same as a non-weighted pair mean technique, except that in the former, the size of each cluster is used as a weight. This technique is particularly useful in an embodiment in which it is suspected that the size of a cluster of transcription control sequences or the like varies considerably (Sneath and Sokal, 1973, "Numerical taxonomy", San Francisco: W.H. Freeman & Co.). Other cluster linkage methods, such as, for example, non-weighted and weighted pair group centroid and Ward's method, are also useful in several embodiments of the present invention. See, for example, Ward, 1963, J. Am. Stat. Assn., 58: 236; and Hartigan, 1975, "Clustering algorithms", New York: Wiley.

In a certain preferred embodiment, cluster analysis can be performed using a well-known hclust technique (e.g., see a well-known procedure in "hclust" available from Program S-Plus, MathSoft, Inc., Cambridge, MA).

According to the present invention, it was found that even if the versatility of stimuli to a clustering set is increased, a state of a cell can be substantially

elucidated by analyzing typically at least two, preferably at least 3, profiles using a method of the present invention. Stimulation conditions include treatment with a pharmaceutical agent in different concentrations, different measurement times after treatment, response to genetic mutations in various genes, a combination of treatment of a pharmaceutical agent and mutation, and changes in growth conditions (temperature, density, calcium concentration, etc.).

10

As used herein, the term "significantly different" in relation to two statistics means that the two statistics are different from each other with a statistical significance. In an embodiment of the present invention, data of a set of experiments assessing the responses of cellular components can be randomized by a Monte Carlo method to define an objective test.

15

In a certain embodiment, an objective test can be defined by the following technique.  $p_{ki}$  represents a response of a component  $k$  in experiment  $i$ .  $\Pi_{(i)}$  represents a random permutation of the indices of experiments. Next,  $p_{k\Pi(i)}$  is calculated for a number of different random permutations (about 100 to 1,000). For each branch of the original tree and each permutation:

25

(1) hierarchical clustering is performed using the same algorithm as that which has been used for the original data which is not permuted (in this case, "hclust"); and

30

(2) an improvement  $f$  in classification in total

variance about the transition is made from on

$$f = 1 - \sum D_k^{(1)} / \sum D_k^{(2)}$$

(6)

5

where  $D_k$  is the square of the distance reference (mean) of component  $k$  with respect to the center of a cluster to which component  $k$  belongs. Superscript 1 or 2 indicates the center of all branches or the center of the more preferable  
 10 cluster of the two subclusters. The distance function  $D$  used in this clustering technique has a considerable degree of freedom. In these examples,  $D=1-r$ , where  $r$  is a correlation coefficient of one response with respect to another response of a component appearing in a set of  
 15 experiments (or of the mean cluster response).

Specifically, an objective statistical test can be preferably used to determine the statistical reliability of grouping any clustering methods or algorithms.  
 20 Preferably, similar tests can be applied to both hierarchical and nonhierarchical clustering methods. The compactness of a cluster is quantitatively defined as, for example, the mean of squares of the distances of elements in the cluster from the "mean of the cluster", or more  
 25 preferably, the inverse of the mean of squares of the distances of elements from the mean of the cluster. The mean of a specific cluster is generally defined as the mean of response vectors of all elements in the cluster. However, in a specific embodiment (e.g., the definition of the mean  
 30 of the cluster is doubtful), for example, the absolute values

of normalized or weighted inner products are used to evaluate the distance function of a clustering algorithm (i.e.,  $I=1-|r|$ ). Typically, the above-described definition of the mean may raise a problem in an embodiment in which response  
5 vectors have opposing directions so that the mean of the cluster as defined above is zero. Therefore, in such an embodiment, a different definition is preferably selected for the compactness of a cluster, for example, without limitation, the mean of squares of the distances of all pairs  
10 of elements in a cluster. Alternatively, the compactness of a cluster may be defined as the mean of distances between each element (e.g., a cellular component) of a cluster and another element of the cluster (or more preferably the inverse of the mean distance).

15

Other definitions, which may be used in statistical techniques used in the present invention, are obvious to those skilled in the art.

20

In another embodiment, a profile of the present invention can be analyzed using signal processing techniques. In these signal processing techniques, a correlation function is defined, a correlation coefficient is calculated, an autocorrelation function and a cross-correlation  
25 function are defined, and these functions are weighted where the sum of the weights is equal to 1. Thereby, moving averages can be obtained.

In signal processing, it is important to  
30 consider a time domain and a frequency domain. Rhythm often

plays an important role in dynamic characteristic analysis for natural phenomena, particularly life and organisms. If a certain time function  $f(t)$  satisfies the following condition, the function is called a periodic function:

5 .

$$f(t) = f(t+T).$$

At time 0, the function takes a value of  $f(0)$ . The function takes a value of  $f(0)$  at time  $T$  again after taking various values after time 0. Such a function is called a periodic function. Such a function includes a sine wave.  $T$  is called a period. The function has one cycle per time  $T$ . Alternatively, this feature may be represented by  $1/T$  which means the number of cycles per unit time (cycles/time) without loss of the information. The concept represented by the number of cycles per unit time is called frequency. If the frequency is represented by  $f$ ,  $f$  is represented by:

20

$$f=1/T.$$

Thus, the frequency is an inverse of the time. The time is dealt in a time domain, while the frequency is dealt in a frequency domain. The frequency may be represented in an electrical engineering manner. For example, the frequency is represented by angular measure where one period corresponds to  $360^\circ$  or  $2\pi$  radians. In this case,  $f$  (cycles/ sec) is converted to  $2\pi f$  (radians/sec), which is generally represented by  $\omega$  ( $=2\pi f$ ) and is called angular frequency.

30

Now, a sine wave is compared with a cosine wave. The cosine wave is obtained by translating the sine wave by  $90^\circ$  or  $\pi/2$  radians. The sine wave may be represented by the delayed cosine wave. This time delay is called phase. For example, when a pure cosine wave has a phase of 0, a sine wave has a phase of  $90^\circ$ . When a sine wave is added to a cosine wave, the amplitude of the resultant wave is increased by a factor of  $\sqrt{2}$  and the phase is  $\pi/4$ .

10

In such analysis, Fourier series and frequency analysis may be available. In addition, Fourier transformation, discrete Fourier transformation, and power spectrum may be available. In Fourier expansion, techniques, such as wavelet transformation and the like, may be available. These techniques are well known in the art and are described in, for example, Yukio Shimizu, "Seimei Sisutemu Kaiseki notameno Sugaku [Mathematics for analyzing life systems]", Corona sha, (1999); and Yasuhiro Ishikawa, "Rinsho Igaku notameno Ueburette Kaiseki [Wavelet analysis for clinical medicine]", Igaku Shuppan.

20

(Description of preferred embodiments)

Hereinafter, the present invention will be described by way of embodiments. The embodiments described below are provided only for illustrative purposes. Accordingly, the scope of the present invention is not limited by the embodiments except as by the appended claims.

25

In one aspect, the present invention provides

30



a method for representing a state of a cell. The method comprises the steps of: a) obtaining a time-lapse profile of the cell by time-lapse monitoring of a gene state (e.g., the expression of a gene (transcription, translation, etc.), etc.) associated with at least one gene selected from genes derived from the cell; and b) presenting the time-lapse profile. For example, the profile of the intensity of a signal obtained by monitoring is subjected to interval differentiation, thereby obtaining a function of changes which can be in turn displayed. In this case, preferably, for example a constitutive promoter or the like, which is assumed to be changed, can be used as a reference to obtain a difference, thereby obtaining a time-lapse profile. The present invention is not limited to this.

15

Time-lapse profiles may be displayed using any method, for example, they may be visually displayed using a display device (e.g., an x axis showing time while the y axis shows signal intensity), or alternatively, may be displayed as a table of numerical values. Alternatively, signal intensity may be displayed as optical intensity. Furthermore, profiles may be presented by means of sound.

Preferably, cells are fixed to a solid phase support (e.g., an array, a plate, a microtiter plate, etc.) when they are monitored. Such fixation can be carried out using techniques known in the art or techniques as described herein. Fixation or immobilization of a cell allows systematic investigation thereof.

30

In a preferred embodiment, such a time-lapse profile may be presented in real time. The real time presentation may contain a time lag to some extent if it is performed substantially in real time. A tolerable time lag is, for example, 10 seconds at maximum, and more preferably 1 second at maximum, though the tolerable time lag depends on the required level of real time (simultaneity).

In another aspect, the present invention provides a method for determining a state of a cell. Such determination of the cellular state is achieved by monitoring changes in a transcriptional state of a transcription control factor, which are not conventionally observed. Therefore, the method of the present invention for determining the cellular state allows determination of various states which cannot be conventionally observed. Such a method comprises the steps of: a) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcriptional state associated with at least one biological agent selected from a biological agent group derived from the cell; and b) determining the state of the cell based on the time-lapse profile of the transcription level.

Preferably, cells are fixed to a solid phase support (e.g., an array, a plate, a microtiter plate, etc.) when they are monitored. Such fixation can be carried out using techniques known in the art or techniques as described herein.

In a preferred embodiment, advantageously, the cellular state determination method of the present invention may further comprise correlating the time-lapse profile with the state of the cell before obtaining the time-lapse profile. Alternatively, such correlation information may be provided from known information. Such a correlating step may be performed at every determining step or correlation information may be stored in a database and used as required.

10

In a preferred embodiment, the transcription control sequence may be, without limitation, a promoter, an enhancer, a silencer, another flanking sequence of a structural gene in a genome, and a genomic sequence other than exons. A promoter is preferable. This is because a transcription level can be directly measured, and the state of transcription directly reflects the state of a cell. In a particular embodiment, the transcription control sequences may include constitutive promoters, specific promoters, inducible promoters, and the like.

20

In certain embodiments, any promoter may be used. The present invention is characterized in that any type of promoter can be used. According to the method of the present invention, profiles can be analyzed from a viewpoint of "procession". Therefore, it is possible to determine a state of a cell using any promoter or any set of promoters. Such determination cannot be achieved by conventional techniques. The present invention is highly useful since the present invention achieves what cannot be achieved by

30

conventional techniques.

In a preferred embodiment, at least two biological agents (for example, transcriptional control  
5 sequence) are monitored. By observing at least two biological agents, 80% of the states of a cell can be typically identified. More preferably, at least 3 biological agents are monitored. By observing at least  
10 three biological agents, at least 90% of the states of a cell can be typically identified. In a most preferred embodiment, at least 8 biological agents are monitored. By observing at least 8 biological agents, substantially all of the states of a cell can be typically identified. Thus, although any biological agents are selected, substantially  
15 all of the states of a cell can be determined by selecting and monitoring a small number of biological agents, as described above. This feature has not been conventionally expected. The method of the present invention is simpler, more precise and more accurate than conventional  
20 determination methods in which observation is made at time points and resultant data is statistically processed as heterologous groups.

Therefore, the determination method of the  
25 present invention preferably further comprises arbitrarily selecting at least one biological agent from a biological agent group before monitoring. An important feature of the present invention is such that a biological agent, which does not exhibit specificity when investigated from point  
30 to point, can be used. Further, the present invention

allows accurate reflection of the resultant data to the state of a cell of interest, since data linearly measured data under a consistent environment can be used. Such accurate data cannot be obtained conventionally.

5

In a preferred embodiment, such a time-lapse profile obtained in the present invention may be presented in real time. Alternatively, in the present invention, data may be obtained in a real time manner. The real time presentation may contain a time lag to some extent if it is performed substantially in real time. A tolerable time lag is, for example, 10 seconds at maximum, and more preferably 1 second at maximum, though the tolerable time lag depends on the required level of real time (simultaneity). As used herein, the term "real time" means that the real time presentation may contain a time lag to some extent if it is performed substantially in real time. A tolerable time lag is, for example, 10 seconds at maximum, and more preferably 1 second at maximum, though the tolerable time lag depends on the required level of real time (simultaneity). For example, the level of real time may be preferably 30 seconds at maximum, or even longer in the case of, for example, therapies required for real time diagnosis.

25 In a particular preferable embodiment, states determined by the cellular state determination method of the present invention includes, for example, differentiated states, undifferentiated states, cellular responses to external factors, cell cycles, growth states, and the like. More specifically, such a state includes, for example,

30

without limitation, a response of a cancer cell to an anticancer agent, drug resistance, a response to a biological clock, a differentiated state of a stem cell (e.g., a mesenchymal stem cell, a neural stem cell, etc.), an undifferentiated state of a purified stem cell (e.g., an embryonic stem cell, etc.), a change in cellular morphology, a state of cellular migration, intracellular localization of a molecule, production of a secreted substance, and the like.

10

Therefore, in a preferred embodiment, a cell assessed by the cellular state determination method of the present invention includes, for example, without limitation, a stem cell or a somatic cell, or a mixture thereof. Alternatively, such a cell includes an adherent cell, a suspended cell, a tissue forming cell, and a mixture thereof.

15

In a preferred embodiment, the cellular state determination method of the present invention may be performed upon a cell fixed on a substrate which is a solid phase support. In such a case, the solid phase support is called a chip. When cells are arrayed on the substrate, the substrate is also called an array.

20

In a particularly preferred embodiment of the cellular state determination method of the present invention, advantageously, when a biological agent (for example, a transcription control sequence) used for determination is a nucleic acid molecule, such a nucleic acid molecule may be operably linked to a reporter gene sequence and may be

25

30

transfected into a cell. In this case, the transcription level of the transcription control sequence can be measured as a signal from the reporter gene.

5

Such transfection may be performed in the solid phase or in the liquid phase. For transfection, a technique for increasing the efficiency of introduction of a target substance into a cell may be used. In the present invention, a target substance (e.g., DNA, RNA, a polypeptide, a sugar chain, or a composite substance thereof, etc.), which cannot be substantially introduced into cells under typical conditions, is presented (preferably, contacted) along with an actin-like substance, such as fibronectin, to a cell, thereby making it possible to efficiency introduce the target substance into cells. Therefore, the transfection method comprises the steps of: A) providing a target substance (i.e., DNA comprising a transcription control sequence) and B) providing an actin-like substance (e.g., fibronectin), wherein the order of steps of A) and B) is not particularly limited, and C) contacting the target substance and the actin-like substance with the cell. The target substance and the actin-like substance may be provided together or separately. The actin-like acting substance may be used as described in detail above for the composition of the present invention for increasing the efficiency of introduction of a target substance into a cell. Such a technique can be carried out as appropriate based on the present specification by those skilled in the art. Therefore, the actin-like substance may be used in a manner

which is described in detail above for the composition of the present invention for increasing the efficiency of introduction of a target substance into a cell. Preferably, the actin-like acting substance may be an extracellular  
5 matrix protein (e.g., fibronectin, vitronectin, laminin, etc.) or a variant thereof. More preferably, fibronectin or a variant or fragment thereof may be used.

In one embodiment, a transcription control  
10 sequence used in the present invention may be capable of binding to a transcription factor. Examples of such a transcription factor include, but are not limited to, ISRE, RARE, STAT3, GAS, NFAT, MIC, AP1, SRE, GRE, CRE, NFκB, ERE, TRE, E2F, Rb, p53, and the like. These transcription  
15 factors are commercially available from BD Biosciences Clontech, CA, USA. ISRE is related to STAT1/2. RARE is related to retinoic acid. STAT3 is related to the control of differentiation. GRE is related to the metabolism of sugar. CRE is related to cAMP. TRE is related to thyroid  
20 hormone. E2F is related to cell cycle. p53 is related to G1 check point. Therefore, such information can be used to determine a state of a cell.

In a preferred embodiment, the determination  
25 step of b) of the present invention comprises comparing the phases of the time-lapse profiles. Phases can be calculated by those skilled in the art using general techniques as described herein above and techniques described in the examples below.

30



In another preferred embodiment, the determination step of b) of the present invention comprises calculating a difference between the time-lapse profile of the cell and a control profile. The difference can be  
5 calculated by those skilled in the art using general techniques as described herein above and techniques described in Examples below.

In another preferred embodiment, the  
10 determination step of b) of the present invention comprises a mathematical process selected from the group consisting of signal processing and multivariate analysis. Such a mathematical process can be easily carried out by those skilled in the art based on the description of the present  
15 specification.

In another aspect, the present invention provides a method for correlating an external factor with a cellular response to the external factor. The method  
20 comprises the steps of: a) exposing a plurality of cells to an external factor on a support capable of retaining the cells in a consistent environment; b) monitoring a transcriptional state relating to at least one of a transcriptional factor group present on or within the cells  
25 over time to generate profile data for the cells; and c) correlating the external factor with the profile.

Any external factor to be correlated in the present invention may be used. Such an external factor is  
30 preferably directly or indirectly applicable to a cell. A

method for applying such an external factor is well known in the art, depending on the type of the external factor used. When a substance is used, the substance is dissolved into a solvent, and the resultant solution is added to a  
5 medium containing a cell.

The correlation method of the present invention may utilize the production method of profiles as described hereinabove.

10

A variety of methods can be provided for correlating a foreign agent and a profile in the method of correlation of the present invention. In brief, profiles obtained when a foreign agent is applied to a cell, are  
15 patternized, and if there is little difference between the patternized profiles, it can be inferred that the particular foreign agent has been applied to the cells.

Preferably, a cell may be monitored in an  
20 immobilized state to a solid support such as an array, a plate, a microtiterplate and the like. Such a method for immobilization can be conducted based on any known methodology in the art or the methods described herein.

25 In a preferred embodiment, a correlation method of the present invention may further comprise using at least two external factors to obtain a profile for each external factor. In certain embodiments, at least three, or at least four, more preferably at least ten such foreign agents may  
30 be used but the present invention is not limited thereto.

More preferably, the correlation step may further comprise dividing at least two profiles into categories and classifying the external factors corresponding to the respective profiles into the categories. Such categorization may be readily conducted by those skilled in the art based on the description of the present specification. Such categorization or classification allows correlation and identification of an unknown foreign agent by means of the method of the present invention.

In a preferred embodiment, a transcription control sequence used in the present invention may be, without limitation, a promoter, an enhancer, a silencer, other flanking sequences of structural genes in genomes, and genomic sequences other than exons. A promoter is preferable, since the transcription level can be directly measured.

In a particular embodiment, transcription control sequences used in the present invention may be constitutive promoters, specific promoters, inducible promoters, and the like. The present invention is characterized in that any type of promoter can be used. According to the method of the present invention, profiles can be analyzed from a viewpoint of "process" or "proccession". Therefore, it is possible to determine a state of a cell using any promoter or any set of promoters. Such determination cannot be achieved by conventional techniques. The present invention is highly useful since the present

invention achieves what cannot be achieved by conventional techniques.

5 In a preferred embodiment, at least two transcription control sequences are monitored. By observing at least two transcription control sequences, at least 80% of the states of a cell can be typically identified. More preferably, at least 3 transcription control sequences are monitored. By observing at least three transcription control sequences, at least 90% of the states of a cell can be typically identified. In a most preferred embodiment, at least 8 transcription control sequences are monitored. By observing at least 8 transcription control sequences, substantially all of the states of a cell can be typically identified. Thus, although any transcription control sequences are selected, substantially all of the states of a cell can be determined by selecting and monitoring a small number of transcription control sequences as described above. This feature has not been conventionally expected. The method of the present invention is simpler, more precise and more accurate than conventional determination methods in which observation is made at time points and resultant data is statistically processed as heterologous groups.

25 Therefore, the determination method of the present invention preferably further comprises arbitrarily selecting at least one transcription control sequence from a group of transcription control sequences before monitoring. An important feature of the present invention is such that a transcription control sequence, which does not exhibit

30

specificity when investigated from point to point, can be used.

In a preferred embodiment, such a time-lapse  
5 profile may be presented in real time. The real time presentation may contain a time lag to some extent if it is performed substantially in real time. A tolerable time lag is, for example, 10 seconds at maximum, and more preferably 1 second at maximum, though the tolerable time  
10 lag depends on the required level of real time (simultaneity). For example, in the case of environment measurement requiring real time identification of external factors, the tolerable time lag may be, for example, 1 sec at maximum, 0.1 sec at maximum, or the like. Alternatively, after data  
15 is stored on a storage medium at real time, profiles may be presented corresponding to the data based on the stored data, with some time lag.

In a preferred embodiment, in the correlation  
20 step of c) of the present invention, the phase of the time-lapse profile may be used as information about the time-lapse profile in order to correlate the external factor with the time-lapse profile. The phase is represented by plus or minus depending on the signal intensity at a certain  
25 time. Even using such a simplified method, a cell or an external factor can be identified, thus demonstrating the precision of the method of the present invention.

Preferably, in the method of the present  
30 invention, cells are advantageously cultured on an array.

This is because a number of cells can be simultaneously observed. Preferably, when a cell is immobilized on a solid support such as an array, a salt may be used.

5                   In a preferred embodiment, the step of monitoring the transcription level over time may comprise obtaining image data from the array. This is because image data can be subjected to visual inspection and a human (particularly, a person skilled in the art, such as a medical practitioner or the like) can easily examine image data with  
10 his/her eyes.

                  In a preferred embodiment of the present invention, the step of correlating the external factor with  
15 the time-lapse profile may comprise distinguishing the phases of the time-lapse profiles. As described above, phase is a simple parameter, and its information processing is simple. Thus, cells can be well identified by such simple information processing.

20                   In a preferred embodiment, examples of an external factor to be identified by the method of the present invention include, but are not limited to, a temperature change, a humidity change, an electromagnetic wave, a  
25 potential difference, visible light, infrared light, ultraviolet light, X-ray, a chemical substance, a pressure, a gravity change, a gas partial pressure, an osmotic pressure, and the like. These factors cannot be satisfactorily identified by conventional methods. By using the cell  
30 determination method of the present invention which places

an importance on "procession", an influence of a factor on a cell can be well examined.

5 In a particularly preferred embodiment, an external factor to be identified by the method of the present invention may be a chemical substance. Examples of such a chemical substance include, but are not limited to, biological molecules, chemical compound, media, and the like.

10

Examples of biological molecules include, but are not limited to, nucleic acids, proteins, lipids, sugars, proteolipids, lipoproteins, glycoproteins, proteoglycans, and the like. These biological molecules are known to have an influence on organisms. Unknown biological molecules are also highly likely to have an influence on organisms and are considered to be important targets for study.

20 Particularly preferably, hormones, cytokine, cell adhesion factors, extracellular matrices, receptor agonists, receptor antagonists, and the like, which are expected to have an influence on cells, are used as biological molecules to be investigated.

25 In another aspect, the present invention provides a method for inferring an unidentified external factor given to a cell based on a time-lapse profile of the cell. The method comprises the steps of: a) exposing the cell to a plurality of known external factors; b) obtaining  
30 a time-lapse profile of the cell for each known external

factor by time-lapse monitoring of a transcription level associated with at least one biological agent selected from the group consisting of biological agents derived from the cell; c) correlating the known external factors with the  
5 respective time-lapse profiles; d) exposing the cell to the unidentified external factor; e) obtaining a time-lapse profile of the unidentified external factor by time-lapse monitoring of the transcription level of the selected biological agent; f) determining a profile corresponding  
10 to the time-lapse profile obtained in the step of e) from the time-lapse profiles obtained in the step of b); and g) determining that the unidentified external factor is the known external factor corresponding to the profile determined in the step of f).

15

In the method of the present invention, the step of exposing a cell to external factors can be performed as described above herein or as illustrated in the examples described below. The step of obtaining a time-lapse profile  
20 can be performed as described above herein or as illustrated in the examples described below. The correlation step can be performed as described above herein or as illustrated in the examples described below. After information about all known external factors has been obtained, an  
25 unidentified external factor is similarly monitored. These pieces of information are compared to determine whether or not the unidentified external factor is a known one. If the profile of an unidentified factor fully matches the profile of a known factor, these two factors can be determined as  
30 being identical. Also, if the profile of an unidentified



factor substantially matches the profile of a known factor, these two factors can be determined to be identical. Such determination depends on the information quantity and quality of the known external factor. Such determination  
5 can be easily carried out by those skilled in the art considering various elements.

In another aspect, the present invention provides a method for inferring an unidentified external  
10 factor given to a cell based on a time-lapse profile of the cell. The method comprises: a) providing data relating to a correlation relationship between known external factors and time-lapse profiles of the cell in response to the known external factors, in relation to at least one promoter  
15 selected from promoters present in the cell; b) exposing the cell to the unidentified external factor; c) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with the selected promoter; d) determining a profile corresponding to the  
20 time-lapse profile obtained in the step of c) from the time-lapse profiles obtained in the step of a); and e) determining that the unidentified external factor is the known external factor corresponding to the profile determined in the step of d).

25

Exposure to external factors, profile generation, correlation, and the like can be carried out using techniques as described herein above or as illustrated in the examples below.

30

In another aspect, the present invention provides a system for presenting a state of a cell. The system comprises: a) means for obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one transcription control sequence selected from the group consisting of transcription control sequences derived from the cell; and b) means for presenting the time-lapse profile. An exemplary system configuration is presented in Figure 32.

A configuration of a computer or system for implementing the cellular state presenting method of the present invention is shown in Figure 17. Figure 17 shows an exemplary configuration of a computer 500 for executing the cellular state presenting method of the present invention. An exemplary system configuration is presented in Figure 32.

The computer 500 comprises an input section 501, a CPU 502, an output section 503, a memory 504, and a bus 505. The input section 501, the CPU 502, the output section 503, and the memory 504 are connected via a bus 505. The input section 501 and the output section 503 are connected to an I/O device 506.

An outline of a process for presenting a state of a cell, which is executed by the computer 500, will be described below.

A program for executing the cellular state

presenting method (hereinafter referred to as a "cellular state presenting program") is stored in, for example, the memory **502**. Alternatively, each component of the cellular state presenting program may be stored in any type of recording medium, such as a floppy disk, MO, CD-ROM, CD-R, DVD-ROM, or the like separately or together. Alternatively, the program may be stored in an application server. The cellular state presenting program stored in such a recording medium is loaded via the I/O device **506** (e.g., a disk drive, a network (e.g., the Internet)) to the memory **504** of the computer **500**. The CPU **502** executes the cellular state presenting program, so that the computer **500** functions as a device for performing the cellular state presenting method of the present invention.

15

Information about a cell or the like is input via the input section **501** as well as profile data obtained. Known information may be input as appropriate.

20

The CPU **502** generates display data based on the information about profile data and cells through the input section **501**, and stored the display data into the memory **504**. Thereafter, the CPU **502** may store the information in the memory **504**. Thereafter, the output section **503** outputs a cellular state selected by the CPU **502** as display data. The output data is output through the I/O device **506**.

25

In another aspect, the present invention provides a system for determining a state of a cell. The system comprises: a) means for obtaining a time-lapse

30

profile of the cell by time-lapse monitoring of a transcription level associated with at least one biological agent selected from the group consisting of biological agents derived from the cell; and b) means for determining  
5 the state of the cell based on the time-lapse profile. An exemplary system configuration is presented in Figure 32.

A configuration of a computer or system for implementing the cellular state determining method of the  
10 present invention is shown in Figure 17. Figure 17 shows an exemplary configuration of a computer 500 for executing the cellular state determining method of the present invention. An exemplary system configuration is presented in Figure 32.

15

The computer 500 comprises an input section 501, a CPU 502, an output section 503, a memory 504, and a bus 505. The input section 501, the CPU 502, the output section 503, and the memory 504 are connected via a bus 505. The input  
20 section 501 and the output section 503 are connected to an I/O device 506.

An outline of a process for determining a state of a cell, which is executed by the computer 500, will be  
25 described below.

A program for executing the cellular state determining method (hereinafter referred to as a "cellular state determining program") is stored in, for example, the  
30 memory 502. Alternatively, each component of the cellular

state determining program may be stored in any type of recording medium, such as a floppy disk, MO, CD-ROM, CD-R, DVD-ROM, or the like, separately or together. Alternatively, the program may be stored in an application  
5 server. The cellular state determining program stored in such a recording medium is loaded via the I/O device **506** (e.g., a disk drive, a network (e.g., the Internet)) to the memory **504** of the computer **500**. The CPU **502** executes the cellular state presenting program, so that the computer **500**  
10 functions as a device for performing the cellular state determining method of the present invention.

Information about a cell or the like is input via the input section **501**, as well as profile data obtained.  
15 Known information may be input as appropriate.

The CPU **502** determines a state of a cell based on the information about profile data and cells input through the input section **501**, generates the results as  
20 determination result data, and stores the determination result data in the memory **504**. Thereafter, the CPU **502** may store the information in the memory **504**. Thereafter, the output section **503** outputs a cellular state selected by the CPU **502** as determination result data. The output data is  
25 output through the I/O device **506**.

In another aspect, the present invention provides a system for correlating an external factor with a response of a cell to the external factor. The system  
30 comprises: a) means for exposing the cell to the external

factor; b) means for obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one promoter selected from the group consisting of promoters derived from the cell; and c) means  
5 for correlating the external factor with the time-lapse profile. Such a system can be implemented using a computer as with the above-described systems. An exemplary system configuration is presented in Figure 32.

10 In another aspect, the present invention provides a system for inferring an unidentified external factor given to a cell based on a time-lapse profile. The system comprising: a) means for exposing the cell to a plurality of known external factors; b) means for obtaining  
15 a time-lapse profile of the cell for each known external factor by time-lapse monitoring of a transcription level associated with at least one biological agent selected from the group consisting of biological agents derived from the cell; c) means for correlating the known external factors  
20 with the respective time-lapse profiles; d) means for exposing the cell to the unidentified external factor; e) means for obtaining a time-lapse profile of the unidentified external factor by time-lapse monitoring of the transcription level of the selected transcription  
25 control sequence; f) means for determining a profile corresponding to the time-lapse profile obtained in the means of e) from the time-lapse profiles obtained in the means of b); and g) means for determining that the unidentified external factor is the known external factor  
30 corresponding to the profile determined in the means of f).

Such a system can be implemented using a computer as with the above-described systems. An exemplary system configuration is presented in Figure 32.

5                   In another aspect, the present invention provides a system for inferring an unidentified external factor given to a cell based on a time-lapse profile, comprising: a) means for providing data relating to a correlation relationship between known external factors and  
10 time-lapse profiles of the cell in response to the known external factors, in relation to at least one biological agent selected from biological agents present in the cell; b) means for exposing the cell to the unidentified external factor; c) means for obtaining a time-lapse profile of the  
15 cell by time-lapse monitoring of a transcription level associated with the selected transcription control sequence; d) means for determining a profile corresponding to the time-lapse profile obtained in the means of c) from the time-lapse profiles obtained in the means of a); and  
20 e) determining that the unidentified external factor is the known external factor corresponding to the profile determined in the means of d). Such a system can be implemented using a computer as with the above-described systems. An exemplary system configuration is presented in  
25 Figure 32.

When the present invention is provided in the form of a system as described above, each constituent element thereof can be implemented as with the detailed or preferred  
30 embodiments of the method of the present invention.

Preferred embodiments of such a system can be easily selected by those skilled in the art and can be made or carried out by those skilled in the art based on the present specification. An exemplary system configuration is  
5 presented in Figure 32.

In another aspect, the present invention provides a computer recordable recording medium recording a program for executing a process for presenting a state  
10 of a cell to a computer. The recording medium records at least a program for executing the procedures of: a) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one biological agent selected from the group consisting of  
15 biological agents derived from the cell; and b) presenting the time-lapse profile.

In another aspect, the present invention provides a computer recordable recording medium recording a program for executing a process for determining a state  
20 of a cell to a computer. The recording medium records at least a program for executing the procedures of: a) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one biological agent selected from the group consisting of  
25 biological agents derived from the cell; and b) determining the state of the cell based on the time-lapse profile of the transcription level.

30 In another aspect, the present invention



provides a computer recordable recording medium recording a program for executing a process for correlating an external factor with a response of a cell to the external factor. The recording medium records at least a program for executing  
5 the procedures of: a) exposing the cell to the external factor; b) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one transcription control factor selected from the group consisting of transcription control factors  
10 derived from the cell; and c) correlating the external factor with the time-lapse profile.

In another aspect, the present invention provides a computer recordable recording medium recording  
15 a program for executing a process for inferring an unidentified external factor given to a cell based on a time-lapse profile. The recording medium records at least a program for executing the procedures of: a) exposing the cell to a plurality of known external factors; b) obtaining  
20 a time-lapse profile of the cell for each known external factor by time-lapse monitoring of a transcription level associated with at least one transcription control factor selected from the group consisting of transcription control factors derived from the cell; c) correlating the known  
25 external factors with the respective time-lapse profiles; d) exposing the cell to the unidentified external factor; e) obtaining a time-lapse profile of the unidentified external factor by time-lapse monitoring of the transcription level of the selected transcription control  
30 sequence; f) determining a profile corresponding to the

time-lapse profile obtained in the procedure of e) from the time-lapse profiles obtained in the procedure of b); and g) determining that the unidentified external factor is the known external factor corresponding to the profile  
5 determined in the procedure of f).

In another aspect, the present invention provides a computer recordable recording medium recording a program for executing a process for inferring an  
10 unidentified external factor given to a cell based on a time-lapse profile. The recording medium records at least a program for executing the procedures of: a) providing data relating to a correlation relationship between known external factors and time-lapse profiles of the cell in  
15 response to the known external factors, in relation to at least one transcription control sequence selected from transcription control sequences present in the cell; b) exposing the cell to the unidentified external factor; c) obtaining a time-lapse profile of the cell by time-lapse  
20 monitoring of a transcription level associated with the selected transcription control sequence; d) determining a profile corresponding to the time-lapse profile obtained in the procedure of c) from the time-lapse profiles obtained in the procedure of a); and e) determining that the  
25 unidentified external factor is the known external factor corresponding to the profile determined in the procedure of d).

When the present invention is provided in the  
30 form of a recording medium as described above, each

constituent element thereof can be implemented as with the detailed or preferred embodiments of the method of the present invention. Preferred embodiments of such a recording medium can be easily selected by those skilled  
5 in the art and can be made or carried out by those skilled in the art based on the present specification.

In another aspect, the present invention provides a program for executing a process for presenting  
10 a state of a cell to a computer. The program executes the procedures of: a) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one biological agent selected from the group consisting of biological agents derived from the cell; and  
15 b) presenting the time-lapse profile.

In another aspect, the present invention provides a program for executing a process for determining a state of a cell to a computer. The program executes the  
20 procedures of: a) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one biological agent selected from the group consisting of biological agents derived from the cell; and  
b) determining the state of the cell based on the time-lapse  
25 profile of the transcription level.

In another aspect, the present invention provides a program for executing a process for correlating an external factor with a response of a cell to the external  
30 factor. The program executes the procedures of:

a) exposing the cell to the external factor; b) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one transcription control factor selected from the group  
5 consisting of transcription control factors derived from the cell; and c) correlating the external factor with the time-lapse profile.

In another aspect, the present invention  
10 provides a program for executing a process for inferring an unidentified external factor given to a cell based on a time-lapse profile. The program executes the procedures of: a) exposing the cell to a plurality of known external factors; b) obtaining a time-lapse profile of the cell for  
15 each known external factor by time-lapse monitoring of a transcription level associated with at least one transcription control factor selected from the group consisting of transcription control factors derived from the cell; c) correlating the known external factors with  
20 the respective time-lapse profiles; d) exposing the cell to the unidentified external factor; e) obtaining a time-lapse profile of the unidentified external factor by time-lapse monitoring of the transcription level of the selected transcription control sequence; f) determining a  
25 profile corresponding to the time-lapse profile obtained in the procedure of e) from the time-lapse profiles obtained in the procedure of b); and g) determining that the unidentified external factor is the known external factor corresponding to the profile determined in the procedure  
30 of f).

In another aspect, the present invention provides a program for executing a process for inferring an unidentified external factor given to a cell based on a time-lapse profile. The program executes the procedures of: a) providing data relating to a correlation relationship between known external factors and time-lapse profiles of the cell in response to the known external factors, in relation to at least one transcription control sequence selected from transcription control sequences present in the cell; b) exposing the cell to the unidentified external factor; c) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with the selected transcription control sequence; d) determining a profile corresponding to the time-lapse profile obtained in the procedure of c) from the time-lapse profiles obtained in the procedure of a); and e) determining that the unidentified external factor is the known external factor corresponding to the profile determined in the procedure of d).

When the present invention is provided in the form of a program as described above, each constituent element thereof can be implemented as with the detailed or preferred embodiments of the method of the present invention. Preferred embodiments of such a program can be easily selected by those skilled in the art and can be made or carried out by those skilled in the art based on the present specification. Description formats of such a program are well known to those skilled in the art and include, for

example, the C+ language, and the like.

In another aspect, the present invention provides a method and system for diagnosing a subject. The  
5 diagnosis method comprises the steps of: a) obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one transcription control sequence selected from the group consisting of transcription control sequences derived from  
10 the cell; b) determining the state of the cell based on the time-lapse profile of the transcription level; and c) determining a condition, disorder or disease of a subject based on the state of the cell. The diagnosis method is provided in the form of a system, the system of the present  
15 invention comprises: a) means for obtaining a time-lapse profile of the cell by time-lapse monitoring of a transcription level associated with at least one transcription control sequence selected from the group consisting of transcription control sequences derived from  
20 the cell; b) means for determining the state of the cell based on the time-lapse profile of the transcription level; and c) means for determining a condition, disorder or disease of a subject based on the state of the cell. The present invention is applicable to tailor-made diagnoses  
25 and therapies, such as drug resistance, selection of appropriate anticancer agents, selection of appropriate transplant cells, and the like. Preferably, the diagnosis method of the present invention may be provided as a therapeutic or preventative method comprising the step of  
30 treating a subject with a therapy or preventative method

selected based on the result of diagnosis. In another preferred embodiment, the diagnosis system of the present invention may be provided as a therapeutic or preventative system comprising means for treating a subject with a therapy or preventative method, selected based on the result of diagnosis. An exemplary system configuration is shown in Figure 32.

A configuration of a computer or system for implementing the diagnosis method and system of the present invention is shown in Figure 17. Figure 17 shows an exemplary configuration of a computer 500 for executing the cellular state determining method of the present invention. An exemplary system configuration is shown in Figure 32.

The computer 500 comprises an input section 501, a CPU 502, an output section 503, a memory 504, and a bus 505. The input section 501, the CPU 502, the output section 503, and the memory 504 are connected via a bus 505. The input section 501 and the output section 503 are connected to an I/O device 506.

An outline of a correlation process, which is executed by the computer 500, will be described below.

A program for executing the correlation method and/or selection of treatment or preventative method (hereinafter referred to as a "correlation program" and a "selection program", respectively) is stored in, for example, the memory 502. Alternatively, the correlation

program and the selection program may be stored in any type of recording medium, such as a floppy disk, MO, CD-ROM, CD-R, DVD-ROM, or the like, separately or together.

Alternatively, the programs may be stored in an application  
5 server. The correlation program and the selection program stored in such a recording medium are loaded via the I/O device **506** (e.g., a disk drive, a network (e.g., the Internet)) to the memory **504** of the computer **500**. The CPU **502** executes the correlation program and the selection  
10 program, so that the computer **500** functions as a device for performing the correlation method and the selection method of the present invention.

The result of analysis of a time-lapse profile  
15 (e.g., phase, etc.) and information about a cell or the like are input via the input section **501**. Secondary information about a condition, disorder or disease to be correlated with a time-lapse profile and information about treatment and/or preventative methods may be input as required.

20

The CPU **502** correlates information about a time-lapse profile with a state of a cell or a condition, disorder or disease of a subject and a preventative or therapeutic method as required, based on the information  
25 input through the input section **501**, and stores correlation data into the memory **504**. Thereafter, the CPU **502** may store the information in the memory **504**. Thereafter, the output section **503** outputs information about a state of a cell or a condition, disorder or disease of a subject and a  
30 preventative or therapeutic method as required, which has



been selected by the CPU **502** as diagnostic information. The output data is output through the I/O device **506**.

(Generation of data)

5           In one aspect, the present invention provides a method for generating profile data of information of a cell. The method comprises the steps of: a) providing and fixing the cell to a support; and b) monitoring a biological agent or an aggregation of biological agents on or within  
10 the cell over time to generate data on the profile of the cell. In this aspect, the present invention is characterized in that the cell is fixed to substantially the same site of the support so that information can be continuously (e.g., in a time-lapse manner, etc.) obtained  
15 from the same cell. Thereby, it is possible to monitor a biological agent and an aggregation of biological agents over time. The time-lapse monitoring makes it possible to obtain a profile of a cell and construct a digital cell. To fix a cell to a support, a fixing agent, such as a salt  
20 or the like, may be used for the support in the present invention. A combination of a salt, a complex of a positively charged substance and a negatively charged substance, and a cell may fix the cell to the support. Any salt may be used in the present invention. Examples of such  
25 a salt include, but are not limited to, calcium chloride, sodium hydrogen phosphate, sodium hydrogen carbonate, sodium pyruvate, HEPES, sodium chloride, potassium chloride, magnesium sulfide, iron nitrate, amino acids, vitamins, and the like. Examples of the above-described combination of  
30 a positively charged substance and a negatively charged

substance include, but are not limited to, complexes of a negatively charged substance selected from the group consisting of DNA, RNA, PNA, a polypeptide, a chemical compound, and a complex thereof and a positively charged  
5 substance selected from the group consisting of a cationic polymer, a cationic lipid, a cationic polyamino acid, and a complex thereof. In a preferred embodiment of the present invention, a biological agent of interest may be a nucleic acid molecule or a molecule derived from such a nucleic acid  
10 molecule. This is because most nucleic acid molecules carry genetic information, from which cellular information can be obtained.

In another aspect, the present invention  
15 relates to data obtained by a method comprising the steps of: a) providing and fixing the cell to a support; and b) monitoring a biological agent or an aggregation of biological agents on or within the cell over time to generate data of the profile of the cell. Such data is obtained by  
20 the method which is not conventionally available, and is thus novel. Therefore, the present invention provides a recording medium storing such data.

In another aspect, the present invention  
25 relates to a method for generating profile data of information of a plurality of cells in a consistent environment. The method comprises the steps of: a) providing a plurality of cells on a support which can maintain a consistent environment; and b) monitoring a  
30 biological agent or an aggregation of biological agents on

or within the cells over time to generate profile data for the cells. In this aspect, the present invention is characterized in that profile data or information for a plurality of cells in a consistent environment can be  
5 obtained. Techniques for providing such an environment is also within the scope of the present invention. To provide a consistent environment for a plurality of cells, a fixing agent, such as a salt or the like, may be used for the support in the present invention. A combination of a salt, a complex  
10 of a positively charged substance and a negatively charged substance, and cells may fix the cells to the support. Any salt may be used in the present invention. Examples of such a salt include, but are not limited to, calcium chloride, sodium hydrogen phosphate, sodium hydrogen carbonate,  
15 sodium pyruvate, HEPES, sodium chloride, potassium chloride, magnesium sulfide, iron nitrate, amino acids, vitamins, and the like. Examples of the above-described combination of a positively charged substance and a negatively charged substance include, but are not limited to, complexes of a  
20 negatively charged substance selected from the group consisting of DNA, RNA, PNA, a polypeptide, a chemical compound, and a complex thereof and a positively charged substance selected from the group consisting of a cationic polymer, a cationic lipid, a cationic polyamino acid and  
25 a complex thereof. In a preferred embodiment of the present invention, a biological agent of interest may be a nucleic acid molecule or a molecule derived from such a nucleic acid molecule. This is because most nucleic acid molecules carry genetic information, from which cellular information can  
30 be obtained.

In a preferred embodiment, an actin-like acting substance is preferably provided to the cells in the method of the present invention. The actin-like acting substance  
5 acts on actin within the cells to deform the internal cytoskeleton to facilitate introduction of an external factor into the cells. The presence of such an actin-like acting substance makes it possible to investigate an influence of an external factor of interest on the cells.  
10

In one embodiment, a biological agent targeted by the present invention is at least one factor selected from the group consisting of nucleic acids, proteins, sugar chains, lipids, low molecular weight molecules, and  
15 composite molecules thereof.

In a preferred embodiment, cells targeted by the present invention are preferably cultured for a certain period of time without stimulation before monitoring. This  
20 procedure is performed for the purpose of synchronizing the target cells. The period of time required for synchronization is, for example, advantageously at least one day, more preferably at least two days, even more preferably at least 3 days, and still even more preferably  
25 at least 5 days. It should be noted that as the period of time for culture is increased, the necessity of maintaining the culture conditions increases. In the synchronization procedure, the same medium is preferably supplied to cells. Therefore, the culture medium is preferably consistent or  
30 at least changed in a consistent manner. To achieve this,

a means for causing convection in the medium may be preferably provided and used.

In a more preferred embodiment, a biological  
5 agent provided to a cell in the present invention may  
comprise a nucleic acid molecule encoding a gene. The  
nucleic acid molecule encoding a gene is preferably  
transfected into a cell. Preferably, such a biological  
agent may be provided along with a transfection reagent (gene  
10 introduction reagent). More preferably, the nucleic acid  
molecule encoding a gene may be provided to a cell along  
with a gene introduction reagent and an actin-like acting  
substance. In this case, the cell is preferably provided  
with a complex of a salt, a positively charged substance,  
15 and a negatively charged substance (in this case, a nucleic  
acid molecule and a gene introduction reagent). Thus, the  
cell and the target molecule are fixed on a support. In  
addition, this technique makes it possible to allow separate  
biological agents (e.g., nucleic acid molecules) to be  
20 separately introduced into cells without a partition. As  
substantially no partition is used, a plurality of cells  
can be monitored in substantially a consistent environment.  
Further, different biological agents can be introduced into  
a cell, thereby making it possible to obtain a profile of  
25 a state of the cell affected by the biological agents. Such  
a profile can be stored as data. Such data may be stored  
in a certain standard format, and therefore, can be  
reproduced and compared. Thus, the present invention has  
an effect which is not achieved by conventional biological  
30 assays. Such data, once obtained and stored in such a

standard format, can be extracted and used for various purposes and a number of times. For example, researchers can perform "virtual experiments" to conduct various analyses under the same conditions while taking into  
5 consideration differences in a substantially infinite number of parameters. In addition, since virtual experiments and the results thereof are stored in a raw data format, undergraduate and graduate students, who otherwise spend most of their school life doing laboratory work, can  
10 receive education in data analysis in the true sense. The above-described cellular profile data can be easily standardized, thereby making it possible to do research based on data which may have been obtained by experiments under the same conditions over the world. Such data may be  
15 distributed in a standardized form. Such a standardized form may be readable to typical computers (e.g., computers having a commonly available OS, such as Windows, Mac, UNIX, LINUX, or the like). Data produced in the present invention may include generated cellular profile data, information  
20 about experimental conditions used in data generation, information about cells, information about environments, and the like.

In a preferred embodiment, a profile targeted  
25 by the present invention may include a profile of gene expression, a profile of an apoptotic signal, a profile of a stress signal, a profile of the localization of a molecule (preferably, the molecule is labeled with a fluorescent, phosphorescent, or radioactive substance, or a combination  
30 thereof), a profile of changes in cellular morphology, a

profile of a promoter, a profile of a promoter dependent on a specific pharmaceutical agent (e.g., antibiotics, ligands, toxins, nutrients, vitamins, hormones, cytokines, etc.), a profile of an intermolecular interaction, and the  
5 like. In an embodiment in which the present invention targets a profile of a promoter dependent on a specific pharmaceutical agent, it is preferable that the present invention may further comprise administering the specific pharmaceutical agent.

10

In a preferred embodiment, the present invention may further comprise providing an external stimulus to the cell. Such an external stimulus may or may not be a biological agent. The external factor may be any  
15 factor and includes, without limitation, substances or other elements (e.g., energy, such as ionizing radiation, radiation, light, acoustic waves, and the like).

In one embodiment, an external factor used in  
20 the present invention may be RNAi. RNAi can be used to substantially suppress an arbitrary gene. It is possible to produce RNAi for all existing genes and investigate the effect of RNAi on the genes. RNAi can be created by techniques well known in the art.

25

In another embodiment, an external factor of the present invention may comprise a chemical substance which does not exist in organisms. By providing such a chemical substance which does not exist in organisms, it is possible  
30 to collect a variety of information. Once collected, such

data can be reused. Therefore, assuming that a chemical substance which does not exist in organisms is not substantially available, if data can be obtained once for such a chemical substance in accordance with the present invention, research can continue without worrying about the availability of such a chemical substance.

In one embodiment, an external factor targeted by the present invention may comprise a ligand to a cellular receptor. By analyzing a ligand, it is possible to study various signal transduction pathways. Therefore, in such a case, a profile obtained according to the present invention may be a profile of receptor-ligand interactions.

In a preferred embodiment of the present invention, a profile of cellular morphology may be obtained. In this case, a method of the present invention may further comprise applying a stimulus to a cell which may be selected from the group consisting of overexpression of a gene, underexpression of a gene, knock down of a gene, addition of an external factor, and a change in an environment.

In a preferred embodiment, a profile obtained according to the present invention may be a profile of interactions between molecules present within a cell. Such an intermolecular interaction includes, but is not limited to, interaction between molecules present in a signal transduction pathway, interaction between a receptor and a ligand, interaction between a transcription factor and a transcription factor sequence, and the like.



In another preferred embodiment, a profile obtained according to the present invention may be a profile of interaction between molecules present in a cell. In this case, a method of the present invention may further comprise observing a cell using a technique selected from the group consisting of a two-hybrid method, FRET, and BRET. The two-hybrid method detects intermolecular interaction within a cell. Specifically, this technique is described in, for example, Protein-Protein Interactions, A MOLECULAR CLONING MANUAL, Edited by Erica Golemis, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (this document also describes FRET). FRET is a technique for detecting inter- or intra-molecular resonance energy shift as a fluorescent wavelength, and is described in, for example, Protein-Protein Interactions (supra); and Miyawaki A., Visualization of the spatial and temporal dynamics of intracellular signaling, Dev. Cell, 2003 Mar; 4(3):295-305. BRET is an intermolecular interaction assay system and is described, for example, Boute N., The use of resonance energy transfer in high-throughput screening: BRET versus FRET, Trends Pharmacol Sci., 2002 Aug; 23(8):351-4.

In a preferred embodiment, cells targeted by the present invention are preferably arranged on a support in a pattern of an array. In this case, preferably, a plurality of cells targeted by the present invention may be spaced at intervals of 10 cm at maximum, more preferably 1 cm at maximum, even more preferably 1 mm at maximum, and most preferably 0.1 mm at maximum. The cells need to be spaced

at minimum intervals. Such intervals may be preferably set so that substantially no interaction occurs.

In one embodiment, a profile obtained according to the present invention may or may not be obtained in real time. A real time profile may be advantageous. When simultaneity is important, it is important to obtain a profile in real time. Alternatively, when a profile is intended to be stored, the profile is not necessarily obtained in real time.

In an additional embodiment, the present invention further comprises fixing a cell to a solid phase support. In this case, the cell is fixed to the solid phase support along with a salt, a complex, an actin-like acting substance, or the like.

In one embodiment, data generated according to the present invention may contain information about a profile. In a preferred embodiment, data generated according to the present invention may contain information about conditions for monitoring, information about a cellular state, information about an external factor, information about an environment, and the like.

In a preferred embodiment, at least two biological agents may be preferably monitored in the present invention, more preferably at least 3 biological agents, and even more preferably at least 8 biological agents. Alternatively, all biological agents in a certain specific

category (e.g., all olfactory receptors, all gustatory receptors, etc.) may be preferably monitored.

Alternatively, in another preferred embodiment,  
5 the present invention may further comprise arbitrarily selecting the above-described biological agents.

In a preferred embodiment, a cell targeted by the present invention may be selected from the group  
10 consisting of stem cells and somatic cells.

In one embodiment, a support used in the present invention is preferably a solid phase support. This is because cells are easily fixed to such a support. Such a  
15 solid phase support may be made of any material known in the art. The support may be in the form of a substrate.

In one embodiment of the present invention, the above-described biological agent may be a nucleic acid and  
20 the above-described cell may be transfected with the nucleic acid. By transfecting the cell with the nucleic acid, an influence of the nucleic acid on the cell can be collected in real time or in a standardized storable format into data or a profile. This cannot be achieved by conventional  
25 techniques. In a preferred embodiment, transfection may be performed in solid a phase or in a liquid phase. More preferably, transfection may be advantageously performed in a solid phase. This is because data collection and standardization or normalization can be more easily carried  
30 out.

In a preferred embodiment of the present invention, a profile may be subjected to a process selected from the group consisting of phase comparison, calculation  
5 of a difference from a control profile, signal processing, and multivariate analysis. Data processed in such a manner may fall within the scope of the present invention.

In another aspect, the present invention  
10 provides a method for presenting profile data of information about a plurality of cells in a consistent environment. The method comprises the steps of: a) providing a plurality of cells on a support capable of retaining the cells in a consistent environment; b) monitoring a biological agent  
15 or an aggregation of biological agents on or within the cells over time to generate profile data for the cells; and c) presenting the data.

The above-described support capable of  
20 retaining a plurality of cells in a consistent environment can be achieved as described elsewhere herein. The step of generating data can be performed as described elsewhere herein. The step of presenting data can be performed as described elsewhere herein. Examples of a method of  
25 performing such presentation include, but are not limited to, techniques of using various sensory means, such as visual means, auditory means, olfactory means, tactile means, gustatory means, and the like. Preferably, a visually presentation means may be used. Such visual means include,  
30 without limitation, a computer display and the like.

Preferably, in the presentation method of the present invention, presentation may be performed in real time. Alternatively, stored data may be stored and  
5 presentation may be delayed. When presentation should be performed in real time, data signals may be transferred directly to, for example, a display.

In another aspect, the present invention  
10 provides a method for determining states of cells in a consistent environment. The method comprises the steps of:  
a) providing a plurality of cells on a support capable of retaining the cells in a consistent environment;  
b) monitoring a biological agent or an aggregation of  
15 biological agents on or within the cells over time to generate profile data for the cells; and c) determining the states of the cells based on the data.

The above-described support capable of  
20 retaining a plurality of cells in a consistent environment can be achieved as described elsewhere herein. The step of generating data can be performed as described elsewhere herein. The step of determining the states of the cells may be performed by correlating the generated data with  
25 information about the cells, or comparing the generated data with standard data. In this case, the data may be statistically processed.

Therefore, in a certain embodiment, the present  
30 invention may further comprise correlating a profile

obtained according to the present invention with a state of a cell before obtaining the time-lapse profile. To perform determination smoothly, the cells targeted by the present invention may advantageously include cells whose  
5 states are known. It is possible to store data of cells whose states are known, determination can thus be quickly performed by comparing data between the known cell and unknown cells.

10 During determination, at least two biological agents are preferably present. In this case, the plurality of biological agents may belong to heterologous categories (e.g., proteins and nucleic acids, etc.) or homologous categories.

15 Preferably, the present invention may further comprise arbitrarily selecting a biological agent. Any biological agent can be selected and used to characterize a state of a cell to some extent, and in some cases,  
20 identification is possible. Thus, the present invention has an effect which cannot be expected from conventional techniques.

In the determination method of the present  
25 invention, data may be preferably generated in real time. When data is generated in real time, an unknown substance or state of an unknown cell may be determined in real time.

In the determination method of the present  
30 invention, examples of a state of a target cell include,

but are not limited to, differentiated states, undifferentiated states, cellular responses to external factors, cell cycles, growth states, and the like.

5           A cell targeted by the present invention may be either a stem cell or a somatic cell. Any somatic cell may be used. A cell may be selected by those skilled in the art, depending on the purpose of use of the cell.

10           A solid phase support used in the determination method of the present invention may comprise a substrate. In the present invention, such a substrate can be used as a part of a computer system, so that determination can be automated. An exemplary configuration of such a system is  
15 shown in Figure 32.

          In a preferred embodiment, in the determination method of the present invention, the biological agent may be a nucleic acid molecule, and the cell is transfected with  
20 the nucleic acid molecule. Transfection may be performed on a solid phase support using any material, but preferably a gene introduction agent, more preferably a salt, an actin-like acting substance, or the like. Transfection may be performed in solid phase or in liquid phase, and  
25 preferably in solid phase.

          In a determination method of the present invention, a target biological agent may be capable of binding to another biological agent. By investigating a  
30 biological agent having such a property, a network mechanism

in a cell may be elucidated.

In a determination method of the present invention, the determination step may comprise a  
5 mathematical process selected from the group consisting of comparison of phases of profiles, collection of differences from a control profile, signal processing, and multivariate analysis. Such processing techniques are well known in the art and described in detail herein.

10

In another aspect, the present invention provides a method for correlating an external factor with a cellular response to the external factor. The method comprises the steps of: a) exposing a plurality of cells  
15 to an external factor on a support capable of retaining the cells in a consistent environment; b) monitoring a biological agent or an aggregation of biological agents on or within the cells over time to generate profile data for the cells; and c) correlating the external factor with the  
20 profile. Exposure of the cells to the external factor may be achieved by placing the cells and the external factor into an environment in which the cells are contacted with the external factor. For example, when the cells are fixed on the support, the external factor is added to the support  
25 to achieve exposure. Techniques for generating and correlating data are also well known in the art, and may be used singly or in combination. Preferably, statistical processes are performed to generate statistically significant data and information.

30



In a preferred embodiment, in the correlation method of the present invention, the cells may be fixed on the support. Since the cells are fixed, data can be easily standardized, so that data can be significantly efficiently  
5 processed.

In a preferred embodiment, a correlation method of the present invention may further comprise using at least two external factors to obtain a profile for each external  
10 factor. Techniques for obtaining such a profile are well described herein.

More preferably, the correlation step may further comprise dividing at least two profiles into  
15 categories and classifying the external factors corresponding to the respective profiles into the categories. By categorization, data can be processed in a more standardized manner.

20 In a preferred embodiment, a profile obtained by the present invention may be presented in real time. When data is intended to be stored, data may not be particularly presented in real time.

25 In a preferred embodiment, a cell used in the present invention may be cultured on an array. In such a case, therefore, the cell is preferably covered with medium. Any medium which is commonly used for cells may be used.

30 In a preferred embodiment of the present

invention, the step of monitoring a profile may comprise obtaining image data from the array. Particularly, when a profile contains visual information (e.g., emission of fluorescence due to gene expression), the profile can be  
5 obtained by capturing image data.

In a correlation method of the present invention, the step of correlating an external factor with a profile may comprise distinguishing between phases of the profile.  
10 Distinguishing phases of the profile can be achieved only after the present invention provides time-lapse profiles obtained in a consistent environment.

An external factor targeted by the present  
15 invention may be selected from the group consisting of a temperature change, a humidity change, an electromagnetic wave, a potential difference, visible light, infrared light, ultraviolet light, X-rays, a chemical substance, a pressure, a gravity change, a gas partial pressure, and an osmotic  
20 pressure. Preferably, the chemical substance may be a biological molecule, a chemical compound, or a medium. Examples of such a biological molecule include, but are not limited to, nucleic acid molecules, proteins, lipids, sugars, proteolipids, lipoproteins, glycoproteins, proteoglycans,  
25 and the like. Such a biological molecule may also be, for example, a hormone, a cytokine, a cell adhesion factor, an extracellular matrix, or the like. Alternatively, the chemical substance may be either a receptor agonist or antagonist.

30

In another aspect, the present invention relates to a method for identifying an unidentified external factor given to a cell from a profile of the cell. The method comprises the steps of: a) exposing a cell to a plurality  
5 of known external factors on a support capable of retaining the cell in a consistent environment; b) monitoring a biological agent or an aggregation of biological agents on or within the cell over time to generate a profile of the cell to each of the known external factors and to generate  
10 profile data for the cell; c) correlating each of the known external factors with each of the profiles; d) exposing the cell to an unidentified external factor; e) monitoring a biological agent or an aggregation of biological agents on or within the cell exposed to the external factors over time  
15 to obtain a profile of the cell with respect to the unidentified external factor; f) determining, from the profiles obtained in the step of b), a profile corresponding to the profile obtained the step of e); and g) determining that the unidentified external factor is the known external  
20 factor corresponding to the profile determined in the step of f). Techniques for exposure to external factors, data generation, correlation, exposure to unidentified external factors, and the like are described elsewhere herein and can be selected as appropriate depending on the purpose by  
25 those skilled in the art taking such descriptions into consideration.

In another aspect, the present invention provides a method for identifying an unidentified external  
30 factor given to a cell from a profile of the cell. The method

comprises the steps of: a) providing data relating to a correlation relationship between known external factors and profiles of the cell in response to the known external factors, in relation to a biological agent or an aggregation  
5 of biological agents on or within the cell; b) exposing the cell to the unidentified external factor; c) monitoring the biological agent or the aggregation of the biological agents on or within the cell to obtain a profile of the cell; d) determining, from the profiles provided in the step of  
10 a), a profile corresponding to the profile obtained in the step of c); and e) determining that the unidentified external factor is the known external factor corresponding to the profile determined in the step of d). Techniques for exposure to external factors, data generation, correlation,  
15 exposure to unidentified external factors, and the like are described elsewhere herein and can be selected as appropriate depending on the purpose by those skilled in the art taking such descriptions into consideration.

20 In another aspect, the present invention provides a method for obtaining a profile relating to information for a plurality of cells in a consistent environment. The method comprises the steps of:  
a) providing a plurality of cells on a support capable of  
25 retaining the cells in a consistent environment; and  
b) monitoring a biological agent or an aggregation of biological agents on or within the cell over time to generate a profile of the cells. Techniques for exposure to external factors, data generation, correlation, exposure to  
30 unidentified external factors, and the like are described

elsewhere herein and can be selected as appropriate, depending on the purpose by those skilled in the art taking such descriptions into consideration.

5                   In another aspect, the present invention relates to a recording medium in which data generated by a method for generating cellular profile data of the present invention is stored. Data may be stored in any format. Any recording medium may be used. Examples of such a recording  
10 medium include, but are not limited to, CD-ROMs, flexible disks, CD-Rs, CD-RWs, MOs, mini disks, DVD-ROMs, DVD-Rs, memory sticks, hard disks, and the like. The present invention also relates to a transmission medium in which data generated by a method for generating cellular profile  
15 data of the present invention is stored. Examples of such a transmission medium include, but are not limited to, networks, such as intranets, the Internet, and the like.

                  A recording medium or transmission medium of the  
20 present invention may further contain data relating to at least one piece of information selected from the group consisting of information about conditions for the monitoring step, information about the profile, information about the cellular state, and information about the  
25 biological agent. Data relating to such information may be stored while being linked to one another. Preferably, the data may be advantageously standardized. Standardized data can be distributed on general distribution pathways. The above-described linkage may be constructed for each cell  
30 or for each biological agent, or for both.

In another aspect, the present invention relates to data generated by a method for generating cellular profile data of the present invention. Such data cannot be  
5 generated by conventional techniques and is thus novel.

In another aspect, the present invention provides a system for generating profile data of information for a plurality of cells in a consistent environment. The  
10 system comprises: a) a support capable of retaining a plurality of cells in a consistent environment; b) means for monitoring a biological factor or an aggregation of biological factors on or within the cells over time; and  
15 c) means for generating profile data for the cells from a signal obtained from the monitoring means. The support capable of retaining cells in a consistent environment can be made by those skilled in the art using a technique first provided by the present invention. Such a technique is  
20 attributed to the finding that cells are fixed and arrayed without a partition. Examples of the monitoring means include, but are not limited to, microscopes (e.g., optical microscopes, fluorescence microscopes, phase-contrast microscopes, etc.), electron microscopes, scanners, naked eyes, infrared cameras, confocal/nonconfocal microscopes,  
25 CCD cameras, and the like. An exemplary configuration of such a system is shown in Figure 32.

In a system of the present invention, the system may not necessarily contain cells from the start, but  
30 preferably may contain cells which are advantageously fixed

on a support. In such a case, fixation is preferably standardized. In addition, the cells are fixed and spaced, for example, without limitation, at intervals of 1 mm or the like.

5

In a preferred embodiment, at least one substance selected from the group consisting of salts and actin-like acting substances may be preferably adhered to the support. By adhering cells to the support with a salt  
10 or an actin-like acting substance, or preferably with both, fixation of the cells and/or introduction of a substance into the cells can be enhanced.

Examples of the monitoring means used in the  
15 system of the present invention include, but are not limited to, optical microscopes, fluorescence microscopes, phase-contrast microscopes, reading devices using a laser source, means using surface plasmon resonance (SPR) imaging, electric signals, chemical or biochemical markers singly  
20 or in combination, radiation, confocal microscopes, nonconfocal microscopes, differential interference microscopes, stereoscopic microscopes, video monitors, infrared cameras, and the like. Preferably, a scanner (e.g., a scanner for scanning a surface of a substrate using a white  
25 light source or laser) may be used. The reason a scanner is preferable is that fluorescence can efficiently transmit excited energy and microscopic technology can be easily applied. Further, measurement can be advantageously performed without significant damage to cells. An  
30 exemplary configuration of such a system is shown in

Figure 32.

In another aspect, the present invention provides a system for presenting a profile of information for a plurality of cells in a consistent environment. The system comprises: a) a support capable of retaining a plurality of cells in a consistent environment; b) means for monitoring a biological factor or an aggregation of biological factors on or within the cells over time; c) means for generating profile data for the cells from a signal obtained from the monitoring means; and d) means for presenting the data. The support, the monitoring means, and the data generating means can be made as described elsewhere herein. The means for presenting data can be achieved by techniques well known in the art. Examples of such a data presenting means include, but are not limited to, computer displays, loudspeakers, and the like. An exemplary configuration of such a system is shown in Figure 32.

A presentation system of the present invention may further comprise a plurality of cells, in which the cells are preferably fixed to the support. In such a case, at least one substance selected from the group consisting of salts and actin-like acting substances may be preferably adhered to the support. By adhering cells to the support with a salt or an actin-like acting substance, or preferably with both, fixation of the cells and/or introduction of a substance into the cells can be enhanced.

Any monitoring means may be used. Examples of



the monitoring means include, but are not limited to, optical microscopes; fluorescence microscopes; phase microscopes; reading devices using a laser source; means using surface plasmon resonance (SPR) imaging, electric signals, chemical or biochemical markers singly or in combination; and the like.

Any data presenting means may be used, including, without limitation, displays, loudspeakers, and the like.

10

In another aspect, the present invention provides a system for determining a state of a cell. The system comprises: a) a support capable of retaining a plurality of cells in a consistent environment; b) means for monitoring a biological factor or an aggregation of biological factors on or within the cells over time; c) means for generating data from a signal obtained by the monitoring means; and d) means for extrapolating the state of the cell from the data. The support, the monitoring means, and the data generating means can be made by those skilled in the art as described elsewhere herein. The means for extrapolating a state of a cell from data may be produced and used by techniques well known in the art. For example, measured data can be compared with standard data for known cells to achieve extrapolation. A device storing a program for such extrapolation or a computer capable of executing such a program may be used as the extrapolation means. An exemplary configuration of such a system is shown in Figure 32.

25  
30

In another aspect, the present invention provides a system for correlating an external factor with responses of cells to the external factor. The system comprises: a) a support capable of retaining a plurality  
5 of cells in a consistent environment; b) means for exposing the cell to the external factor; c) means for monitoring a biological factor or an aggregation of biological factors on or within the cells over time; d) generating profile data for the cells from a signal from the monitoring means; and  
10 e) means for correlating the external factor with the profile. The support, the monitoring means, and the data generating means can be made by those skilled in the art as described elsewhere herein. The means for exposing the cells to the external factor can be designed and carried  
15 out as appropriate by those skilled in the art depending on the properties of the external factor. The correlation means can employ a recording medium storing a program for correlation or a computer capable of executing such a program. Preferably, a system of the present invention comprises a  
20 plurality of cells. An exemplary configuration of such a system is shown in Figure 32.

In another aspect, the present invention provides a system for identifying an unidentified external  
25 factor given to a cell based on a profile of the cell. The system comprises: a) a support capable of retaining a plurality of cells in a consistent environment; b) means for exposing the cell to one or more known external factors; c) means for monitoring a biological factor or an  
30 aggregation of biological factors on or within the cells

over time; d) means for obtaining a profile of the cell with respect to each of the known external factors to generate profile data for the cell; e) means for correlating each of the known external factors with each profile; f) means  
5 for exposing the cell to the unidentified external factor; g) means for comparing the profiles of the known external factors obtained by the means of d) with the profile of the unidentified external factor to determine a profile of the  
10 unidentified external factor from the profiles of the known external factors, wherein the determined unidentified external factor is the known external factor corresponding to the determined profile. The support, the exposure means, the monitoring means, the data generating means, and the correlation means, and the other exposure means can be made  
15 and carried out as appropriate by those skilled in the art as described elsewhere herein. The means for determining a corresponding profile can also be made and carried out by utilizing a recording medium storing a program capable of executing such a determination process and a computer  
20 capable of executing such a program. Preferably, a system of the present invention comprises a plurality of cells. An exemplary configuration of such a system is shown in Figure 32.

25 In another aspect, the present invention provides a system for identifying an unidentified external factor given to a cell based on a profile of the cell. The system comprises: a) a recording medium storing providing data relating to a correlation relationship between known  
30 external factors and profiles of the cell in response to

the known external factors, in relation to a biological factor or an aggregation of biological factors on or within the cell; b) means for exposing the cell to the unidentified external factor; c) a support capable of retaining a plurality of cells in a consistent environment ; d) means for monitoring a biological factor or an aggregation of biological factors on or within the cells over time; e) means for obtaining a profile of the cell from a signal obtained by the monitoring means; f) means for determining, from the profiles stored in the recording medium of a), a profile corresponding to the profile obtained with respect to the unidentified external factor, wherein the determined unidentified external factor is the known external factor corresponding to the determined profile. The support, the exposure means, the monitoring means, the data generating means, and the correlation means, and the other exposure means can be made and carried out as appropriate by those skilled in the art as described elsewhere herein. The means for determining a corresponding profile can also be made and carried out by utilizing a recording medium storing a program capable of executing such a determination process and a computer capable of executing such a program. Preferably, a system of the present invention comprises a plurality of cells. An exemplary configuration of such a system is shown in Figure 32.

In another aspect, the present invention relates to a support capable of maintaining a consistent environment for a plurality of cells. Such a support was first provided by the present invention. By utilizing such

a support, a plurality of cells can be analyzed in a consistent environment.

Preferably, cells are arranged on a support in the form of an array. This is because standardized analysis can be achieved. In this case, the support may preferably comprise a salt or an actin-like acting substance. More preferably, the support may advantageously comprise a complex of a positively charged substance and a negatively charged substance. This is because cells can be easily fixed to the support using such a complex. Actin-like acting substances are preferable when the interior of cells is analyzed, since the actin-like acting substances increase the efficiency of introduction of external factors into cells. Therefore, in a preferred embodiment of the present invention, the support may comprise a salt and an actin-like acting substance, and more preferably may comprise a complex of a positively charged substance and a negatively charged substance.

20

A support of the present invention is characterized in that cells may be provided and spaced at intervals of 1 mm. In the case of such intervals, it is not conventionally possible to provide an environment without a partition. Therefore, the present invention has a remarkable effect, as well as practicability, applicability and utility.

In a preferred embodiment, a support of the present invention may comprise a cell fixed thereto. In a more preferred embodiment, a support of the present

30

invention may comprise a biological factor fixed thereto.

In a preferred embodiment, at least two biological factors may be fixed to the support. Such  
5 biological factors may be factors selected from the group consisting of nucleic acid molecules, proteins, sugars, lipids, metabolites, low molecular weight molecules, and complexes thereof, and factors containing physical elements and/or temporal elements.

10

In a more preferred embodiment, a cell and a biological factor may be fixed to a support of the present invention in a mixed manner. The biological factor and the cell may be provided so that they can interact with each  
15 other. Such interaction may vary depending on the biological factor. According to the properties of the biological factor, those skilled in the art can understand how the biological factor interacts with the cell and where the biological factor is positioned so as to interact with  
20 the cell.

In a preferred embodiment, a salt, a complex of a positively charged substance and a negatively charged substance, and an actin-like acting substance are fixed  
25 along with a cell and a biological factor to a support of the present invention.

In a more preferred embodiment, a salt, a complex of a positively charged substance and a negatively  
30 charged substance, and an actin-like acting substance are

fixed along with a cell and a biological factor to a support of the present invention in the form of an array. With such a structure, a cell chip capable of generating the profile data of a cell can be provided. The support has a structure  
5 in which a salt, a complex of a positively charged substance and a negatively charged substance, and an actin-like acting substance are fixed along with a cell and a biological factor in the form of an array. Such a support is also called a "transfection array".

10

Examples of a salt used in the support of the present invention include, but are not limited to, calcium chloride, sodium hydrogen phosphate, sodium hydrogen carbonate, sodium pyruvate, HEPES, sodium chloride,  
15 potassium chloride, magnesium sulfide, iron nitrate, amino acids, vitamins, and the like. A preferable salt is, for example, without limitation, sodium chloride or the like.

Examples of a gene introduction agent used in  
20 the support of the present invention include, but are not limited to, cationic polymers, cationic lipids, polyamine-based reagents, polyimine-based reagents, calcium phosphate, oligofectamin, and oligofectors and the like. Preferably the gene introduction reagents used may  
25 be preferably, but are not limited to lipofectamines, oligofectamines and oligofectors.

Examples of an actin-like acting substance used in the support of the present invention include, but are  
30 not limited to, fibronectin, laminin, vitronectin, and the

like. A preferable actin-like acting substance is, for example, without limitation, fibronectin.

Examples of a nucleic acid molecule used in the support of the present invention include, but are not limited to, nucleic acid molecules comprising transcription control sequences (e.g., promoters, enhancers, etc.), gene coding sequences, genomic sequences containing nontranslation regions, nucleic acid sequences encoded by the genome of a host (a fluorescent protein gene, E. coli/yeast self-replication origins, a GAL4 domain, etc.), and the like. Preferable nucleic acid molecules include, but are not limited to, transcription control sequences (e.g., promoters, enhancers, etc.), gene coding sequences, genomic sequences containing nontranslation regions, and the like.

Examples of a cell used in the support of the present invention include, but are not limited to, stem cells, established cell lines, primary culture cells, insect cells, bacterial cells, and the like. Preferable cells include, but are not limited to, stem cells, established cell lines, primary culture cells, and the like.

Examples of a material for a support of the present invention include, but are not limited to, glass, silica, plastics, and the like. Preferable materials include, but are not limited to, the above-described materials with coating.

In another aspect, the present invention



provides a method for producing a support comprising a plurality of cells fixed thereto and capable of maintaining a consistent environment for the cells. The method comprises the steps of: A) providing the support; and  
5 B) fixing the cells via a salt and a complex of a positively charged substance and a negatively charged substance onto the support. The step of providing a support may be achieved by obtaining a commercially available support or molding a support material. A support material may be prepared by  
10 mixing starting materials for the material as required. The fixing step can be carried out by using techniques known in the art. Examples of such fixing techniques include, but are not limited to, an ink jet printing technique, a pin array technique, a stamping technique, and the like. These  
15 techniques are well known and can be performed as appropriate by those skilled in the art.

In a preferred embodiment, the fixing step in the present invention may comprise fixing a mixture of the  
20 salt, the complex of a gene introduction agent and an actin-like acting substance (positively charged substances) and a nucleic acid molecule (a negatively charged substance), and the cell in the form of an array. Such a fixing step may be achieved by printing techniques.  
25

In another aspect, the present invention provides a device for producing a support comprising a plurality of cells fixed thereto and capable of maintaining  
30 a consistent environment for the cells. The device

comprises: A) means for providing the support; and B) means for fixing the cells via a salt and a complex of a positively charged substance and a negatively charged substance onto the support. The support may be obtained using means which  
5 can perform the above-described methods. Examples of such means include, but are not limited to, a support molding means, a material formulating means (e.g., a mixing means), and the like. The molding means can employ techniques well known in the art. The fixing means may comprise a printing  
10 means. As such a printing means, commercially available ink jet printers can be used.

(Digital Cell)

As used herein the term "digital cell" refers to a  
15 collection of at least one experimental data corresponding to a cell of experimental interest. Such experimental data is a correlation between the conditions used for the experiments conducted on an actual cell in the real world, and the experimental results thereof. The digital cell is  
20 composed such that when an experimental condition is given, an experimental result relating to the experimental condition will be reproduced.

By using the digital cell, experimental results  
25 conducted on an actual cell can be reproduced on a computer system. This allows institutions or individuals having no experimental facilities to conduct cutting-edge studies relating to a cell. As a result, it allows the introduction of business entities having a primary interest in  
30 disciplines other than that of the present technical art,

which could not been achieved to date, prior to the disclosure of the present invention.

Figure **33A** depicts an example of data structure of a digital cell. This example represents a digital cell by a collection of three experimental data A1, A2 and A3 relating to cell A.

Each of experimental data A1, A2 and A3, comprises cell parameter, environment parameter and stimulus parameter as parameters indicating experimental conditions, and stimulus response result as an experimental result.

As used herein, the cell parameter specifies a cell of experimental interest. The environment parameter specifies an environment under which the cell specified by the cell parameter is cultured. The stimulus parameter specifies a stimulus given to the cell specified by the cell parameter. The stimulus response result shows a result which the cell specified by the cell parameter responded to the stimulus specified by the stimulus parameter under the environment specified by the environment parameter.

Experimental data A1 shows that cell A was cultured in a medium called "DMEM", under the culture condition of pH "7", temperature "37" degree Celsius, CO<sub>2</sub> concentration "5" %, and a stimulus consisting of a reporter called "Tet-OFF CMV EGF" or "MCV EGFP" and a chemical stimulus (agent) "Doxycyclene" is given thereto to obtain

a stimulus response result. The stimulus response result is represented by "cell dynamic data 1" and "reporter measurement data 1".

5                   Experimental data A2 shows that cell A was cultured in a medium called "DMEM", under the culture condition of pH "7", temperature "37" degree Celsius, CO<sub>2</sub> concentration "5" %, and a stimulus consisting of a reporter called "c-fos" and a chemical stimulus (agent) "PSC833" is  
10 given thereto to obtain a stimulus response result. The stimulus response result is represented by "cell dynamic data 2" and "reporter measurement data 2".

                  Experimental data A3 shows that cell A was  
15 cultured in a medium called "DMEM", under the culture condition of pH "5", temperature "39" degree Celsius, CO<sub>2</sub> concentration "4" %, and a stimulus consisting of a reporter called "CREB" and a chemical stimulus (agent) "Vindesine" is given thereto to obtain a stimulus response result. The  
20 stimulus response result is represented by "cell dynamic data 3" and "reporter measurement data 3".

                  As such, parameters indicating experimental conditions (a cell parameter, an environment parameter and  
25 a stimulus parameter) and a stimulus response result showing an experimental result are correlated. Such correlation and data correlated thereby are called experimental data. The digital cell is provided as a collection of at least one experimental data on a cell of experimental interest.

30

Figure **33B** shows another example of data structure of the digital cell. This example shows the layered structure of the data structure shown in Figure **33A**. As such, layering the structure of the data structure of the digital cell allows expression of the same content with  
5 less data than the data structure shown in Figure **33A**.

In the examples of Figures 33A and 33B, correlation has been presented by a unidirectional link  
10 (arrows in the Figures) between the parameter showing the experimental conditions and experimental results. However, methods of correlation are not limited thereto. Any methods of correlation may be used herein.

15 (Production of a digital cell)

Figure **34** shows an example of the procedure of a process of producing a digital cell. This procedure may be implemented by any type of computer.

20 Step S3401: Cell parameters specifying a cell of experimental interest are obtained. Cell parameters can be obtained by, for example, receiving cell parameters inputted by a user into a computer. Alternatively, data outputted from an experimental apparatus may be obtained by collection  
25 or analysis of the same in an automatic manner by a computer to obtain cell parameters.

Step S3402: Environment parameters specifying an environment under which the cell specified by the cell  
30 parameters is cultured, are obtained. The environment

parameters are obtained by receiving, by a computer, environment parameters inputted by a user, for example. Alternatively, environment parameters may be obtained by automatically collecting or analyzing data outputted from an experimental apparatus (for example, sensors measuring experimental environment and the like) and the like, by a computer. The environment parameters include, for example, parameters representing a medium for culturing a cell and a parameter representing conditions for such culture. Parameters of such culture conditions include for example, pH, temperature, CO<sub>2</sub> concentration of the medium, and the like.

Step S3403: Stimulus parameter specifying a stimulus to be given to a cell specified by the cell parameter. A stimulus parameter is obtained by, for example, receiving, by a computer, a stimulus parameter inputted by a user. Alternatively, a stimulus parameter may be obtained by automatically collecting or analyzing, by a computer, data outputted by an experimental apparatus. Such a stimulus parameter may comprise, for example, a parameter representing a reporter and a parameter representing a chemical stimulus.

Step S3404: A stimulus response result showing the result in response to a stimulus by the stimulus parameter, by a cell specified by the cell parameter under the environment specified by the environment parameter, is obtained. The stimulus response result is obtained by automatically collecting or analyzing data outputted from

an experimental apparatus such as monitoring apparatus for monitoring the course of experiments.

Step S3405: The cell parameter, the environment  
5 parameter, the stimulus parameter and the stimulus response  
result are correlated with each other. This correlation  
allows production of an experimental data against a cell  
of experimental interest. Such a correlation is conducted  
by linking in a single direction shown in Figure **33A**.  
10 However, methods correlation is not limited to such.

Step S3406: Steps S3401 through S3405 are  
repeated as necessary. This allows production of at least  
one experimental data against a cell of experimental  
15 interest. The collection of at least one experimental data  
is provided as a digital cell. .

The computer implementing the process for  
producing a digital cell, functions as an apparatus or device  
20 for producing a digital cell. The digital cell produced is  
stored on, for example, a database which can be accessed  
by the computer.

As such, provision of a digital cell of a  
25 collection of at least one experimental data, is only  
possible by the present inventors by providing and  
developing technologies for locating a plurality of cells  
on a substrate under a consistent environment.  
Conventionally, in the prior art, it was not possible to  
30 maintain a plurality of cells under a consistent environment,

and thus the experimental conditions have not been reliable,  
and thus no significance is found when accumulating  
experimental data between experiments. As such, the  
"production of a digital cell" is a real advance in  
5 technology which is feasible for the first time through the  
technological innovation of the present inventors.

(Provision of a method of reproducing  
experimental results against an actual cell)

10 Figure 35 depicts an example of a configuration  
of computer system 3501 which provides a service reproducing  
an experimental result obtained using an actual cell using  
the digital cell.

15 Computer system 3501 comprises service  
requester 3510 requesting a service desired by a user, and  
service provider 3520 providing the desired service in  
response to the request.

20 Computer system 3501 may comprise a plurality of  
service requesters 3510.

Service provider 3520 is configured so as to be  
capable of accessing database 3522 with at least one digital  
25 cell stored thereon. A database structure of the digital  
cell stored on database 3522 is shown in, for example,  
Figures **33A** and **33B**. Database 3522 may be provided inside  
service provider 3520, or may be located outside service  
provider 3520.

30



Service provider 3520 may be configured so as to be capable of accessing a plurality of databases stored thereon with respect to at least one digital cell.

5           Service requester 3510 and service provider 3520 may independently be any type of computer.

Service requester 3510 and service provider 3520 are connected to each other via network 3530. Network  
10 3530 may be any type of network, but in view of feasibility of connection or cost, most preferably, the network is the Internet.

When network 3530 is the Internet, service  
15 requester 3510 may be a Web browser operated by a user, and service provider 3520 may be a Web server connected to service requester via the Internet. Such a configuration allows any user across the entire world easy access to service provider 3520.

20

Figure **36** depicts an example of process for providing a service of reproducing an experimental result against an actual cell using a digital cell. This process may be implemented by cooperating service requester 3510  
25 and service provider 3520.

Step S3601: Service requester 3510 receives cell and environment parameters and produces request comprising the cell parameter, the environment parameter and the stimulus  
30 parameter. The request is described in, for example, XML.

Step S3602: Service request 3510 provides the request to service provider 3520.

5       Step S3603: Service provider 3520 searches for database 3522 in response to the request, to determine whether or not there is a stimulus response result relating to the cell parameter, the environment parameter and the stimulus parameter included in the request in database 3522.

10

Step S3604: when it is determined that there exists a stimulus response result relating to the cell parameter, the environment parameter and the stimulus parameter included in the request in database 3522, service provider 15 3520 provides service requester 3510 with the stimulus response result. The stimulus response result is described in, for example, XML.

Step S3605: Service requester 3510 displays the 20 stimulus response result provided by service provider 3520.

If it is determined there is no stimulus response result relating to the cell parameter, the environment parameter and the stimulus parameter included in the 25 requested in database 3522, service provider 3520 provides service requester 3510 with a result of "no hit", for example.

Procedures as shown in Figure 36 may be processed in 30 a single computer. For example, a single computer program

in a single computer may be used for implementing the procedures of steps S3601 through S3605 shown in Figure 36. In this case, such a single computer functions as an apparatus having the combined functions of service requester  
5 3510 and service provider 3520.

Figure 37 depicts an example of input interface for inputting a cell parameter, an environment parameter and a stimulus parameter to service 3510. In this example,  
10 these parameters are inputted by inputting these parameters as texts by a user into a desired region.

Any number of methods may be employed as a method for inputting these parameters into service requester 3510.  
15 For example, these parameters may be inputted by choosing these parameters from a menu (such as, pull-down menu, pop-up menu and the like) by a user.

Service requester 3510 may employ any  
20 embodiment for displaying the stimulus response result. For example, service requester 3510 may display the stimulus response result on a display screen, or may output the stimulus response result to a printer. Service requester 3510 may display the stimulus response result using a still  
25 image or display the stimulus response result using movie display.

The stimulus response result may include profile data of a cell obtained by monitoring a biological  
30 agent or a collection thereof on or in a cell over time.

In such a case, for example, the profile data of a cell shown in Figure 19 may be displayed by service requester 3510 as the stimulus response result.

5           As such, according to computer system 3510, it is now possible to provide a service of reproducing an experimental result for an actual cell using the digital cell. As such, it is possible to conduct an advanced search relating to a cell even by a research organization or an  
10 individual without experimental facilities.

Figure 38 depicts an example of configurations of computer system 3801 for providing a service of reproducing an experimental result against an actual cell  
15 using the digital cell.

Computer system 3801 comprises service requester 3810 requesting a service desired by a user; a plurality of service providers 3820<sub>1</sub>, 3820<sub>N</sub>; and service registry  
20 3840 with registration of at least one service which can be provided by a plurality of service providers 3820<sub>1</sub>, 3820<sub>N</sub>, wherein N is an any integer of two or more.

Computer system 3801 may include a plurality of  
25 service requesters 3810. Service provider 3820<sub>i</sub> is configured so as to be capable of accessing database 3822<sub>i</sub> at least one digital cell stored thereon. A data structure of a digital cell stored on database 3822<sub>i</sub> is as shown in Figures 33A and 33B. Database 3822<sub>i</sub> may be provided in  
30 service provider 3820<sub>i</sub> or outside service provider 3820,

wherein  $i = 1, 2, \dots, N$ .

Service provider 3820<sub>i</sub> may be configured to be capable of accessing a plurality of databases with at least one  
5 digital cell each stored thereon.

Service registry 3840 is configured to be capable of accessing database 3842 with data stored thereon representing services being capable of being provided by  
10 service providers 3820<sub>i</sub> to 3820<sub>N</sub>. Database 3842 may be provided in service registry or outside service registry 3840. Storing data representing services on database 3842 allows registration of services to service registry 3840. Formats of data stored on database 3842 are preferably  
15 pre-normalized. Storage of data to database 3842 may be performed manually by a firm managing service registry 3840 or by transmitting data from service providers 3820<sub>i</sub> to 3820<sub>N</sub> via network 3830 to service registry 3840.

20 Each service requester 3810, service provider 3820<sub>i</sub> to 3820<sub>N</sub> and service registry 3840 may be any type of computer.

Each of service providers 3820<sub>i</sub> to 3820<sub>N</sub> is preferably  
25 conducted by research carried out by any of organizations, firms or any other corporation possessing experimental facilities which conducts research on an actual cell. Each of service requester 3810 and service registry 3840 is preferably conducted by any of organizations, firms or any  
30 other corporation (for example, an association for promoting

digital cells) managing provision of services of reproducing experimental results against an actual cell using the digital cell. Further, in order to secure the quality of services registered to service registry 3840, it is  
5 preferable to oblige such an organization which manages service providers  $3820_i$  to  $3820_N$  to satisfy a predetermined standard.

Service requester 3810, service provider  $3820_i$   
10 to  $3820_N$  and service registry 3840 are connected via network 3830. Network 3830 is of any type but most preferably, in view of ease of connection and cost, is the Internet.

When network 3830 is the Internet, service  
15 requester 3810 may be a Web server connected to a Web browser operated by a user via the Internet. Each of  $3820_i$ ,  $3820_N$  may be a Web server connected to service requester 3810 via the Internet. In this case, service requester 3810 functions as portal or Website interrelaying to a Web browser operated  
20 by a user and a Web server of service provider  $3820_i$ . This configuration allows easy access to service providers of  $3820_i$  to  $3820_N$  by users all over the world, and it is now possible for research institutes and/or firms all over the world to participate in the business of providing the service  
25 of reproducing experimental results against an actual cell using a digital cell.

Figure 39 depicts an example of a .process for providing a service of reproducing an experimental result  
30 against an actual cell using the digital cell. This

procedure is implemented by cooperating service requester 3810 and service providers 3820<sub>i</sub> with 3820<sub>N</sub>.

Step S3910: Service requester 3810 receives a  
5 cell parameter, an environment parameter, and a stimulus parameter, and produces a request comprising such a cell parameter, an environment parameter, and a stimulus parameter. Such a request is described in, for example, XML.

10

Step S3902: Service requester 3810 searches service registry 3840 responding to the request, and determines whether or not there exists a service provider 3820<sub>i</sub> which can provide a service of the requester amongst  
15 service providers 3820<sub>i</sub> to 3820<sub>N</sub>, wherein *i* is any integer of 1 to *N*.

Service providers 3820<sub>i</sub> to 3820<sub>N</sub> may employ any type of method for register services which can be provided by  
20 service providers 3820<sub>i</sub> to 3820<sub>N</sub> on service registry 3840. For example, when service provider 3820<sub>1</sub> is capable of providing a service of reproducing an experimental result against cell A, then cell parameters specifying cell A and addresses (for example, URL and the like) specifying the  
25 locations of service provider 3820<sub>1</sub> may be stored on database 3842. For example, if service provider 3820<sub>2</sub> can provide services of reproducing cells B and C, then cell parameters specifying cell A and addresses (for example, URL and the like) specifying the locations of service provider 3820<sub>2</sub> may  
30 be stored on database 3842. Alternatively, when service

provider can provide the service of reproducing experimental results satisfying specific experimental conditions against cell D, then parameters such as environment parameters and stimulus parameters specifying the experimental conditions and addresses (for example, a URL and the like) specifying the locations of service provider 3820<sub>3</sub> may be stored on database 3842.

Step S3903: If there exists a service provider 3820<sub>i</sub> which can provide a service of the requester, amongst service providers 3820<sub>1</sub> to 3820<sub>N</sub>, such service requester 3810 provides service provider 3820<sub>i</sub> with the request. The location of service provider 3820<sub>i</sub> may be specified by referring to database 3842 of service registry 3840.

Step S3904: service provider 3820<sub>i</sub> searches database 3822<sub>i</sub> in response to the request, and determines whether or not there exists the stimulus response result relating to the cell parameter, the environment parameter and the stimulus parameter included in the request in database 3822<sub>i</sub>.

Step S3905: If determined that there exists a stimulus response result relating to the cell parameter, the environment parameter and the stimulus parameter included in the request in database 3822<sub>i</sub>, service provider 3820<sub>i</sub> provides service requester 3810 with the stimulus response result. The stimulus response result is described in, for example, XML.



Step S3906: service requester 3810 displays stimulus response result provided by service provider 3820<sub>i</sub>.

5 If determined that there is no stimulus response result relating to the cell parameter, environment parameter and stimulus parameter contained in the request in database 3822<sub>i</sub>, service provider 3820<sub>i</sub> will provide service requester 3810 with, for example, the result of "no hit".

10 As described above, any number of methodologies may be employed as a method for inputting a cell parameter, an environment parameter and a stimulus parameter to service requester 3810, and further any forms may be employed as a form of displaying stimulus response result by service  
15 requester 3810.

As such, according to computer system 3810, it is possible to provide a service of reproducing an experimental result against an actual cell using the digital  
20 cell. This allows research institutes or individual having no experimental facilities to perform advanced research activities relating to a cell. Further, according to computer system 3801, registration of services capable of being provided by a plurality of service providers 3820<sub>i</sub> to  
25 3820<sub>N</sub>, provides opportunities to participate in the business of providing the service of reproducing experimental results against an actual cell using the digital cell to research organizations or firms all over the world.

30 All patents, published patent applications and

publications cited herein are incorporated by reference as if set forth fully herein.

The preferred embodiments of the present invention have been heretofore described for a better understanding of the present invention. Hereinafter, the present invention will be described by way of examples. Examples described below are provided only for illustrative purposes. Accordingly, the scope of the present invention is not limited except as by the appended claims. According to the examples below, it will be understood that those skilled in the art can select cells, supports, biological agents, salts, positively charged substances, negatively charged substances, actin-like acting substances, and the like, as appropriate, and can make or carry out the present invention.

#### EXAMPLES

Hereinafter, the present invention will be described in greater detail by way of examples, though the present invention is not limited to the examples below. Reagents, supports, and the like are commercially available from Sigma (St. Louis, USA), Wako Pure Chemical Industries (Osaka, Japan), Matsunami Glass (Kishiwada, Japan) unless otherwise specified.

(Example 1: Reagents)

Formulations below were prepared in Example 1.

30

As candidates for an actin-like acting substance, various extracellular matrix proteins and variants or fragments thereof were prepared in Example 1, as listed below. Fibronectin and the like were commercially available. Fragments and variants were obtained by genetic engineering techniques:

- 1) fibronectin (SEQ ID NO.: 11);
- 2) fibronectin 29 kDa fragment;
- 10 3) fibronectin 43 kDa fragment;
- 4) fibronectin 72 kDa fragment;
- 5) fibronectin variant (SEQ ID NO.: 11, alanine at 152 was substituted with leucine);
- 6) ProNectin F (Sanyo Chemical Industries, Kyoto, Japan);
- 15 7) ProNectin L (Sanyo Chemical Industries);
- 8) ProNectin Plus (Sanyo Chemical Industries);
- 9) laminin (SEQ ID NO.: 6);
- 10) RGD peptide (tripeptide);
- 11) RGD-containing 30kDa peptide;
- 20 12) 5 amino acids of laminin (IKVAV, SEQ ID NO.: 28); and
- 13) gelatin.

Plasmids were prepared as DNA for transfection. Plasmids, pEGFP-N1 and pDsRed2-N1 (both from BD Biosciences, Clontech, CA, USA) were used. In these plasmids, gene expression was under the control of cytomegalovirus (CMV). The plasmid DNA was amplified in E. coli (XL1 blue, Stratgene, TX, USA) and the amplified plasmid DNA was used as a complex partner. The DNA was dissolved in distilled water free from DNase and RNase.

The following transfection reagents were used:  
Effectene Transfection Reagent (cat. no. 301425, Qiagen,  
CA), TransFast™ Transfection Reagent (E2431, Promega, WI),  
5 Tfx™-20 Reagent (E2391, Promega, WI), SuperFect  
Transfection Reagent (301305, Qiagen, CA), PolyFect  
Transfection Reagent (301105, Qiagen, CA), LipofectAMINE  
2000 Reagent (11668-019, Invitrogen corporation, CA),  
JetPEI (×4) conc. (101-30, Polyplus-transfection, France),  
10 and ExGen 500 (R0511, Fermentas Inc., MD). These  
transfection reagents were added to the above-described DNA  
and actin-like acting substances in advance, or complexes  
thereof with the DNA were produced in advance.

15 The thus-obtained solution was used in assays  
using transfection arrays described below.

(Example 2: Transfection array - Demonstration  
using mesenchymal stem cells)

20 In Example 2, an improvement in the  
transfection efficiency in solid phase was observed. The  
protocol used in Example 2 will be described below.

(Protocol)

25 The final concentration of DNA was adjusted to  
1 µg/µL. An actin-like acting substance was preserved as  
a stock having a concentration of 10 µg/µL in ddH<sub>2</sub>O. All  
dilutions were made using PBS, ddH<sub>2</sub>O, or Dulbecco's MEM. A  
series of dilutions, for example, 0.2 µg/µL, 0.27 µg/µL,  
30 0.4 µg/µL, 0.53 µg/µL, 0.6 µg/µL, 0.8 µg/µL, 1.0 µg/µL,

1.07  $\mu\text{g}/\mu\text{L}$ , 1.33  $\mu\text{g}/\mu\text{L}$ , and the like, were formulated.

Transfection reagents were used in accordance with instructions provided by each manufacturer.

5

Plasmid DNA was removed from a glycerol stock and amplified in 100 mL L-amp overnight. Qiaprep Miniprep or Qiagen Plasmid Purification Maxi was used to purify DNA in accordance with a standard protocol provided by the  
10 manufacturer.

In Example 2, the following 5 cells were used to confirm an effect: human mesenchymal stem cell (hMSCs, PT-2501, Cambrex BioScience Walkersville, Inc., MD); human  
15 embryonic renal cell (HEK293, RCB1637, RIKEN Cell Bank, JPN); NIH3T3-3 cell (RCB0150, RIKEN Cell Bank, JPN); HeLa cell (RCB0007, RIKEN Cell Bank, JPN); and HepG2 (RCB1648, RIKEN Cell Bank, JPN). These cells were cultured in DMEM/10% IFS containing L-glut and pen/strep.

20

(Dilution and DNA spots)

Transfection reagents and DNA were mixed to form a DNA-transfection reagent complex. Formation of the complex requires a certain period of time. Therefore, the  
25 mixture was spotted onto a solid phase support (e.g., a poly-L-lysine slide) using an arrayer. In Example 2, as a solid phase support, an APS slide, a MAS slide, and an uncoated slide were used as well as a poly-L-lysine slide. These slides are available from Matsunami Glass (Kishiwada,  
30 Japan) or the like.

For complex formation and spot fixation, the slides were dried overnight in a vacuum dryer. Drying was performed for a duration in the range of 2 hours to 1 week.

5

Although the actin-like acting substance might be used during the complex formation, it was also used immediately before spotting in Example 2.

10 (Formulation of mixed solution and application to solid phase supports)

300  $\mu$ L of DNA concentrated buffer (EC buffer) + 16  $\mu$ L of an enhancer were mixed in an Eppendorf tube. The mixture was mixed with a Vortex, followed by incubation for  
15 5 minutes. 50  $\mu$ L of a transfection reagent (Effectene, etc.) was added to the mixture, followed by mixing by pipetting. To apply a transfection reagent, an annular wax barrier was formed around the spots on the slide. 366  $\mu$ L of the mixture was added to the spot region surrounded by  
20 the wax, followed by incubation at room temperature for 10 to 20 minutes. Thereby, the fixation to the support was manually achieved.

(Distribution of cells)

25 Next, a protocol for adding cells will be described. Cells were distributed for transfection. The distribution was typically performed by reduced-pressure suction in a hood. A slide was placed on a dish, and a solution containing cells was added to the dish for  
30 transfection. The cells were distributed as follows.

The growing cells were distributed to a concentration of  $10^7$  cells/25 mL. The cells were plated on the slide in a 100×100×15 mm squared Petri dish or a 100 mm (radius) × 15 mm circular dish. Transfection was conducted for about 40 hours. This period of time corresponded to about 2 cell cycles. The slide was treated for immunofluorescence.

10 (Evaluation of gene introduction)

Gene introduction was evaluated by detection using, for example, immunofluorescence, fluorescence microscope examination, laser scanning, radioactive labels, and sensitive films, or emulsion.

15

When an expressed protein to be visualized is a fluorescent protein, such a protein can be observed with a fluorescence microscope and a photograph thereof can be taken. For large-sized expression arrays, slides may be scanned using a laser scanner for storage of data. If an expressed protein can be detected using fluorescent antibodies, an immunofluorescence protocol can be successively performed. If detection is based on radioactivity, the slide may be adhered as described above, and autoradiography using film or emulsion can be performed to detect radioactivity.

25

(Laser scanning and Quantification of fluorescence intensity)

30 To quantify transfection efficiency, the

present inventors use a DNA microarray scanner (GeneTAC UC4x4, Genomic Solutions Inc., MI). Total fluorescence intensity (arbitrary unit) was measured, and thereafter, fluorescence intensity per unit surface area was calculated.

5

(Cross-sectional observation by confocal scanning microscope)

Cells were seeded on tissue culture dishes at a final concentration of  $1 \times 10^5$  cells/well and cultured in appropriate medium (Human Mesenchymal Cell Basal Medium (MSCGM BulletKit PT-3001, Cambrex BioScience Walkersville, Inc., MD). After fixation of the cell layer with 4% paraformaldehyde solution, SYTO and Texas Red-X phalloidin (Molecular Probes Inc., OR, USA) was added to the cell layer for observation of nuclei and F-actin. The samples emitting light due to gene products and the stained samples were observed with a confocal laser microscope (LSM510: Carl Zeiss Co., Ltd., pinhole size=Ch1=123  $\mu\text{m}$ , Ch2=108  $\mu\text{m}$ , image interval = 0.4) to obtain cross sectional views.

20

(Results)

Figure 1 shows the results of experiments in which various actin-like acting substances and HEK293 cells were used where gelatin was used as a control.

25

As can be seen from the results, whereas transfection was not very successful in a system using gelatin, transfection took place to a significant level in systems using fibronectin, ProNectin (ProNectin F, ProNectin L, ProNectin Plus) which is a variant of

30



fibronectin, and laminin. Therefore, it was demonstrated that these molecules significantly increased transfection efficiency. Use of the RGD peptide alone exhibited substantially no effect.

5

Figures 2 and 3 show transfection efficiency when fibronectin fragments were used. Figure 4 shows the summary of the results. 29 kDa and 72 kDa fragments exhibited a significant level of transfection activity, while a 43 kDa fragment had activity but its level was low. Therefore, it was suggested that an amino acid sequence contained in the 29 kDa fragment played a role in an increase in transfection efficiency. Substantially no contamination was found in the case of the 29 kDa fragment, while contamination was observed in the case of the other two fragments (43 kDa and 72 kDa). Therefore, only the 29 kDa domain may be preferably used as an actin-like acting substance. When only the RGD peptide was used, increased transfection efficiency was not exhibited. The 29-kDa peptide therefore exhibited activity with respect to enhancing transfection efficiency. Such a system with an additional 6 amino acids of laminin (higher molecular weight) exhibited transfection activity. Therefore, these peptide sequences may also play an important role in increased transfection efficiency, without limitation. In such a case, a molecular weight of at least 5 kDa, preferably at least 10 kDa, and more preferably at least 15 kDa may be required for an increase in transfection efficiency.

30

Next, Figure 5 shows the result of studies on

the transfection efficiency of cells. In Figure 5, HEK293 cells, HeLa cells, and 3T3 cells, which were conventionally transfectable, and HepG2 cells and mesenchymal stem cells (MSC) which were conventionally believed to be substantially impossible to transfect, were used to show the effect of the transfection method of the present invention. The vertical axis represents the intensity of GFP.

In Figure 5, the transfection method of the present invention using a solid phase support was compared with a conventional liquid phase transfection method. The conventional liquid phase transfection method was conducted in accordance with a protocol recommended by the kit manufacturer.

As can be seen from Figure 5, transfection efficiency comparable to HeLa and 3T3 was achieved in HepG2 cells and mesenchymal stem cells (MSC) which were conventionally believed to be substantially impossible to transfect, as well as HEK293 cells, HeLa cells, and 3T3 cells, which were conventionally transfectable. Such an effect was not achieved by conventional transfection systems. The present invention was the first to provide a system which can increase transfection efficiency for substantially all cells and can provide practicable transfection to all cells. By using solid phase conditions, cross contamination was significantly reduced. Therefore, it was demonstrated that the present invention using a solid phase support is appropriate for production of an integrated bioarray.

Next, Figure 6 shows the results of transfection when various plates were used. As can be seen from the results of Figure 6, when coating was provided, contamination was reduced as compared with when coating was not provided and transfection efficiency was increased.

Next, Figure 7 shows the results of transfection where the concentration of fibronectin was 0, 0.27, 0.53, 0.8, 1.07, and 1.33 ( $\mu\text{g}/\mu\text{L}$  for each). In Figure 7, slides coated with PLL (poly-L-lysine), APS and uncoated slides were shown.

As can be seen from the results of Figure 7, transfection efficiency was increased with an increase in fibronectin concentration. Note that in the case of PLL coating and the absence of coating, the transfection efficiency reached a plateau at a fibronectin concentration of more than 0.53  $\mu\text{g}/\mu\text{L}$ . In the case of APS, it was found that the effect was further increased at a fibronectin concentration of more than of 1.07  $\mu\text{g}/\mu\text{L}$ .

Next, Figure 8 shows photographs indicating cell adhesion profiles in the presence or absence of fibronectin. Figure 9 shows cross-sectional photographs. It was revealed that the morphology of adherent cells were significantly different (Figure 8). The full extension of cells was found for the initial 3 hours of culture in the presence of fibronectin, while extension was limited in the absence of fibronectin (Figure 9). Considering the behavior of filaments (Figure 9) and the results of the

time-lapse observation, it was considered that an actin-like acting substance, such as fibronectin, attached to a solid phase support had an influence on the shape and orientation of actin filaments, and the efficiency of introduction of a substance into a cell, such as transfection efficiency or the like, was thus increased. Specifically, actin filaments quickly change their location in the presence of fibronectin, and disappear from the cytoplasmic space under the nucleus as the cell extends. It is considered that actin depletion in the perinuclear space, which is induced by an actin-like acting substance, such as fibronectin, allows the transport of a target substance, such as DNA or the like, into cells or nuclei. Though not wishing to be bound by any theory, the reason is considered to be that the viscosity of cytoplasm is reduced and positively charged DNA particles are prevented from being trapped by negatively charged actin filaments. Additionally, it is considered that the surface area of the nucleus is significantly increased in the presence of fibronectin (Figure 10), possibly facilitating the transfer of a target substance, such as DNA or the like, into nuclei.

(Example 3: Application to bioarrays)

Next, larger-scale experiments were conducted to determine whether or not the above-described effect was demonstrated when arrays were used.

(Experimental protocols)

(Cell sources, culture media, and culture conditions)

In this example, five different cell lines were used: human mesenchymal stem cells (hMSCs, PT-2501, Cambrex BioScience Walkersville, Inc., MD), human embryonic kidney cell HEK293 (RCB1637, RIKEN Cell Bank, JPN), NIH3T3-3  
5 (RCB0150, RIKEN Cell Bank, JPN), HeLa (RCB0007, RIKEN Cell Bank, JPN), and HepG2 (RCB1648, RIKEN Cell Bank, JPN). In the case of human MSCs, cells were maintained in commercialized Human Mesenchymal Cell Basal Medium (MSCGM BulletKit PT-3001, Cambrex BioScience Walkersville, Inc.,  
10 MD). In case of HEK293, NIH3T3-3, HeLa and HepG2, cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM, high glucose 4.5 g/L with L-Glutamine and sodium pyruvate; 14246-25, Nakalai Tesque, JPN) with 10% fetal bovine serum (FBS, 29-167-54, Lot No. 2025F, Dainippon Pharmaceutical  
15 CO., LTD., JPN). All cells were cultivated in a controlled incubator at 37°C in 5% CO<sub>2</sub>. In experiments involving hMSCs, we used hMSCs of less than five passages, in order to avoid phenotypic changes.

20 (Plasmids and Transfection reagents)

To evaluate the efficiency of transfection, the pEGFP-N1 and pDsRed2-N1 vectors (cat. no. 6085-1, 6973-1, BD Biosciences Clontech, CA) were used. Both genes' expressions were under the control of cytomegalovirus (CMV)  
25 promoter. Transfected cells continuously expressed EGFP or DsRed2, respectively. Plasmid DNAs were amplified using Escherichia coli, XL1-blue strain (200249, Stratagene, TX), and purified by EndoFree Plasmid Kit (EndoFree Plasmid Maxi Kit 12362, QIAGEN, CA). In all cases, plasmid DNA was  
30 dissolved in DNase and RNase free water. Transfection

reagents were obtained as below: Effectene Transfection Reagent (cat. no.301425, Qiagen, CA), TransFast™ Transfection Reagent (E2431, Promega, WI), Tfx™-20 Reagent (E2391, Promega, WI), SuperFect Transfection Reagent (301305, Qiagen, CA), PolyFect Transfection Reagent (301105, Qiagen, CA), LipofectAMINE 2000 Reagent (11668-019, Invitrogen corporation, CA), JetPEI (×4) conc. (101-30, Polyplus-transfection, France), and ExGen 500 (R0511, Fermentas Inc., MD).

10

(Solid-Phase Transfection Array (SPTA) production)

The detail of protocols for 'reverse transfection' are described in the web site, 'Reverse Transfection Homepage' (http://staffa.wi.mit.edu/sabatini\_public/reverse\_transfection.htm) or J. Ziauddin, D. M. Sabatini, Nature, 411, 2001, 107; and R.W. Zu, S.N. Bailey, D.M. Sabatini, Trends in Cell Biology, Vol. 12, No. 10, 485. In our solid phase transfection (SPTA method), three types of glass slides were studied (silanized glass slides; APS slides, and poly-L-lysine coated glass slides; PLL slides, and MAS coated slides; Matsunami Glass, JPN) with a 48 square pattern (3 mm × 3 mm) separated by a hydrophobic fluoride resin coating.

25

(Plasmid DNA printing solution preparation)

Two different ways to produce a SPTA were developed. The main differences reside in the preparation of the plasmid DNA printing solution.

30

## (Method A)

In the case of using Effectene Transfection Reagent, the printing solution contained plasmid DNA and cell adhesion molecules (bovine plasma fibronectin (cat. no. 16042-41, Nakalai Tesque, JPN), dissolved in ultra-pure water at a concentration of 4 mg/mL). The above solution was applied on the surface of the slide using an inkjet printer (synQUAD<sup>TM</sup>, Cartesian Technologies, Inc., CA) or manually, using a 0.5 to 10  $\mu$ L tip. This printed slide was dried up over 15 minutes at room temperature in a safety-cabinet. Before transfection, total Effectene reagent was gently poured on the DNA-printed glass slide and incubated for 15 minutes at room temperature. The excess Effectene solution was removed from the glass slide using a vacuum aspirator and dried up at room temperature for 15 minutes in a safety-cabinet. The DNA-printed glass slide obtained was set in the bottom of a 100-mm culture dish and approximately 25 mL of cell suspension (2 to  $4 \times 10^4$  cells/mL) was gently poured into the dish. Then, the dish was transferred to the incubator at 37°C in 5% CO<sub>2</sub> and incubated for 2 or 3 days.

## (Method B)

In case of other transfection reagents (TransFast<sup>TM</sup>, Tfx<sup>TM</sup>-20, SuperFect, PolyFect, LipofectAMINE 2000, JetPEI ( $\times 4$ ) conc., or ExGen), plasmid DNA, fibronectin, and the transfection reagent were mixed homogeneously in a 1.5-mL micro-tube according to the ratios indicated in the manufacturer's instructions and incubated at room

temperature for 15 minutes before printing on a chip. The printing solution was applied onto the surface of the glass-slide using an inkjet printer or a 0.5- to 10- $\mu$ L tip. The printed glass-slide was completely dried up at room temperature over 10 minutes in a safety-cabinet. The printed glass-slide was placed in the bottom of a 100-mm culture dish and approximately 3 mL of cell suspension (2 to  $4 \times 10^4$  cells/mL) was added and incubated at room temperature over 15 minutes in a safety-cabinet. After incubation, fresh medium was poured gently into the dish. Then, the dish was transferred to an incubator at 37°C in 5% CO<sub>2</sub> and incubated for 2 to 3 days. After incubation, using fluorescence microscopy (IX-71, Olympus PROMARKETING, INC., JPN), we observed the transfectants, based on their expression of enhanced fluorescent proteins (EFP, EGFP and DsRed2). Phase contrast images were taken with the same microscope. In both protocols, cells were fixed using a paraformaldehyde (PFA) fixation method (4% PFA in PBS, treatment time was 10 minutes at room temperature).

20

(Laser scanning and fluorescence intensity quantification)

In order to quantify the transfection efficiency, we used a DNA micro-array scanner (GeneTAC UC4x4, Genomic Solutions Inc., MI). The total fluorescence intensity (arbitrary units) was measured, and thereafter, the fluorescence intensity per surface area was calculated.

25

(Results)

30

(Fibronectin-supported

localized



transfection)

A transfection array chip was constructed as shown in Figure 11. The transfection array chip was constructed by microprinting a cell cultivation medium  
5 solution containing fibronectin and DNA/transfection reagent onto a poly L lysine (PLL) coated glass slide.

Various cells were used for this example. The cells were cultivated under typical cell cultivation  
10 conditions. As they adhered to the glass slide, the cells efficiently incorporated and expressed the genes corresponding to the DNA printed at a given position on the array. As compared to conventional transfection methods (e.g., cationic lipid or cationic polymer-mediated  
15 transfection), the efficiency of transfection using the method of the present invention was high in all the cells tested. Importantly, it was found that tissue stem cells, such as HepG2 and hMSC, which were conventionally believed to resist transfection, were efficiently transfected. hMSC  
20 was transfected at an efficiency 40 or more times higher than that of conventional techniques. In addition, high spatial localization, which is required for high-density arrays, was achieved (low cross contamination between adjacent spots on the array). This was confirmed by  
25 production of a checkered pattern array of EGFP and Ds-Red. hMSC cultivated on this array expressed the corresponding fluorescent proteins with virtually total space resolution. The result is shown in Figure 12. As can be seen from Figure 12, it was found that there was little cross  
30 contamination. Based on the study of the role of the

individual components of the printed mixture, transfection efficiency can be optimized.

5 (Solid-phase transfection array of human mesenchymal stem cells)

The capacity of human Mesenchymal Stem Cells (hMSC) to differentiate into various kinds of cells is particularly intriguing in studies which target tissue regeneration and renewal. In particular, the genetic  
10 analysis of transformation of these cells has attracted attention with expectation of understanding of a factor that controls the pluripotency of hMSC. In conventional hMSC studies, it is not possible to perform transfection with desired genetic materials.

15

(Solid-phase transfection array of human mesenchymal stem cells)

The capacity of human Mesenchymal Stem Cells (hMSC) to differentiate into various kinds of cells is particularly intriguing in studies which target tissue  
20 regeneration and renewal. In particular, the genetic analysis of transformation of these cells has attracted attention with expectation of understanding a factor that controls the pluripotency of hMSC. In conventional hMSC  
25 studies, it is not possible to perform transfection with desired genetic materials.

To achieve this, conventional methods include either a viral vector technique or electroporation. The  
30 present inventors developed a complex-salt system, which

could be used to achieve solid phase transfection which makes it possible to obtain high transfection efficiency to various cell lines (including hMSC) and special localization in high-density arrays. An outline of solid phase  
5 transfection is shown in Figure 13A.

It was demonstrated that solid phase transfection can be used to achieve a "transfection patch" capable of being used for *in vivo* gene delivery and a solid  
10 phase transfection array (SPTA) for high-throughput genetic function research on hMSC.

Although a number of standard techniques are available for transfecting mammalian cells, it is known that  
15 it is inconvenient and difficult to introduce genetic material into hMSC as compared with cell lines, such as HEK293, HeLa, and the like. Conventional viral vector delivery and electroporation techniques are both important. However, these techniques have the following  
20 inconveniences: potential toxicity (for the virus technique); difficulty in high-throughput analysis at the genomic scale; and limited applications *in vivo* studies (for electroporation).

25 The present inventors developed a solid phase support fixed system which can be easily fixed to a solid phase support and has sustained-release capability and cell affinity, whereby most of the above-described drawbacks could be overcome.

30

An example of the results of the above-described experiment is shown in Figure **13B**. The present inventors used our microprinting technique to fix a mixture of a selected genetic material, a transfection reagent, an appropriate cell adhesion molecule, and a salt onto a solid support. By culturing cells on a support having such a mixture fixed thereonto, the gene contained in the mixture was taken in by the cultured cells. As a result, it became possible to facilitate support-adherent cells to take in DNA spatially separated therefrom (Figure **13B**).

As a result of this example, several important effects were achieved: high transfection efficiency (thereby making it possible to study a group of cells on a statistically significant scale); low cross contamination between regions having different DNA molecules (thereby making it possible to study the effects of different genes separately); the extended survival of transfected cells; high-throughput, compatible and simple detecting procedure. SPTA having these features serves as an appropriate basis for further studies.

To achieve the above-described objects, the present inventors studied five different cell lines (HEK293, HeLa, NIH3T3, HepG2 and hMSC) as described above with both our methodology (transfection in a solid phase system) (see Figures **13A** and **13C**) and conventional liquid-phase transfection under a series of transfection conditions. Cross contamination was evaluated for both systems as follows. In the case of SPTA, we printed DNA's encoding a

red fluorescent protein (RFP) and a green fluorescent protein (GFP) on glass supports in a checked pattern. In the case of experiments including conventional liquid phase transfection (where cells to be transfected cannot be spatially separated from one another spontaneously), a DNA encoding GFP was used. Several transfection reagents were evaluated: four liquid transfection reagents (Effectene, TransFast™, Tfx™-20, LopofectAMINE 2000), two polyamine (SuperFect, PolyFect), and two polyimine (JetPEI (x4) and ExGen 500).

Transfection efficiency: transfection efficiency was determined as total fluorescence intensity per unit area (Figure 14A and Figure 14B (images)). The results of liquid phase optimal for cell lines used were obtained using different transfection reagents (see Figures 14C to 14D). Next, these efficient transfection reagents were used to optimize a solid phase protocol. Several tendencies were observed. For cell lines which are readily transfectable (e.g., HEK293, HeLa, NIH3T3, etc.), the transfection efficiency observed in the solid phase protocol was slightly superior to, but essentially similar to, that of the standard liquid phase protocol (Figure 14A to 14D).

However, for cells which are difficult to transfect (e.g., hMSC, HepG2, etc.), we observed that transfection efficiency was increased up to 40 fold while the features of the cells were retained under conditions optimized to the SPTA methodology (see the above-described

protocol and Figures **14C** and **14D**). In the case of hMSC (Figures **15A** and **15B**), the best conditions included use of a polyethylene imine (PEI) transfection reagent. As expected, important factors for achieving high transfection efficiency are the charge balance (N/P ratio) between the number of nitrogen atoms (N) in the polymer and the number of phosphate residues (P) in plasmid DNA, and DNA concentration. Generally, increases in the N/P ratio and the concentration lead to an increase in transfection efficiency. We also observed a significant reduction in the survival rate of hMSC cells in liquid phase transfection experiments where the DNA concentration was high and the N/P ratio was high. Because of these two opposing factors, the liquid phase transfection of hMSC had a relatively low cell survival rate (N/P ratio >10). In the case of the SPTA protocol, however, a considerably high N/P ratio (fixed to the solid support) and DNA concentration were tolerable (probably attributable to the effect of the solid support stabilizing cell membranes) while the cell survival rate and the cellular state were not significantly affected. Therefore, this is probably responsible for the dramatic improvement in transfection efficiency. It was found that the N/P ratio of 10 was optimal for SPTA, and a sufficient transfection level was provided while minimizing cytotoxicity. Another reason for the increase in transfection efficiency observed in the case of the SPTA protocol is that a high local ratio of the DNA concentration to the transfection reagent concentration was achieved (this leads to cell death in liquid phase transfection experiments).

The coating agent used is crucial to achieving of high transfection efficiency on chips. It was found that when a glass chip is used, PLL provided best results both  
5 for transfection efficiency and cross contamination (described below). When fibronectin coating was not used, few transfectants were observed (all the other experimental conditions were retained unchanged). Although it's function is not completely established, fibronectin  
10 probably plays a role in accelerating the cell adhesion process (data not shown), and thus limits the time which permits the diffusion of DNA released from the surface.

Low cross contamination: apart from the higher  
15 transfection efficiency observed in the SPTA protocol, an important advantage of the technique of the present invention is the provision of an array of separated cells, in which selected genes are expressed in the separate positions. The present inventors printed JetPEI (see the  
20 "Experimental protocols" section) and two different reporter genes (RFP and GFP) mixed with fibronectin on glass surface coated with fibronectin. The resultant transfection chip was subjected to appropriate cell culture. Expressed GFP and RFP were localized in regions in which  
25 corresponding cDNA had been spotted, under experimental conditions which had been found to be best. Substantially no cross contamination was observed (Figures **16A** to **16D**). In the absence of fibronectin or PLL, however, cross contamination which hinders solid phase transfection was  
30 observed, and the transfection efficiency was significantly

lower (see Figure 6). This result demonstrated the hypothesis that the relative proportion of plasmid DNA, which was released from the cell adhesion and the support surface, is an important factor in high transfection efficiency and high cross contamination.

Another cause of cross contamination may be the mobility of transfected cells on a solid support. The present inventors measured both the rate of cell adhesion (Figure 16C) and the diffusion rate of plasmid DNA on several supports. As a result, substantially no DNA diffusion occurred under optimum conditions. However, a considerable amount of plasmid DNA diffused under high cross contamination conditions until cell adhesion was completed, so that plasmid DNA was depleted from the solid phase surface.

This established technique is of particular importance in the context of cost-effective high-throughput gene function screening. Indeed, the small amounts of transfection reagent and DNA required, as well as the possible automatization of the entire process (from plasmid isolation to detection) increase the utility of the above presented method.

25

In conclusion, the present invention has successfully realized a hMSC transfection array in a system using complex-salt. With this technique, it will be possible to achieve high-throughput studies using solid phase transfection, such as the elucidation of the genetic

30



mechanism underpinning the differentiation of pluripotent stem cells. The detailed mechanism of the solid phase transfection as well as methodologies for the use of this technology for high throughput, real time gene expression  
5 monitoring can be applied for various purposes.

(Example 4: Mathematical analysis)

Next, time-lapse profiles were produced based on data obtained using the techniques described in  
10 Examples 2 and 3.

(Induction of differentiation)

Each reporter was fixed to a solid phase support and cultured in undifferentiated mesenchymal stem cell maintenance medium (MSCGM, PT-3001, PT-3238, PT-4105, Cambrex, BioWhittaker, USA) for two days. Thereafter, the medium was replaced with differentiation-inducing medium (hMSC Differentiation, PT-3002, PT-4120, Cambrex, BioWhittaker, USA). The response profile of each reporter  
15  
20 was measured.

(Mathematical analysis technique)

A mathematical analysis technique used herein is shown in Figures **18A** and **18B (18-1 to 18-2)**.  
25

(Transcription factors used herein)

As shown in Figures **19** and **24**, plasmids (commercially available from Clontech), in which 17 transcription factors (ISRE, RARE, STAT3, GAS, NFAT, MIC, AP1, SRE, GRE, CRE, NFkB, ERE, TRE, E2F, Rb, p53) were operably  
30

linked to GFP, were used to observe the differentiation of mesenchymal stem cells into osteoblasts. The resultant time-lapse profiles are shown in Figure 19. Reporters for the transcription factors were constructed as shown in  
5 Figure 23.

An assay was conducted using the reporters for the transcription factors under control conditions (cells, supplemental factors, culture conditions, etc.) published  
10 by Clontech.

The results are shown in Figure 25. It was demonstrated that when compared only to DNA in this manner, most of the transcription factors were induced when inducing  
15 agents were added.

Next, the activity of the transcription factors was measured over time in the course of induction of differentiation into bone. In this case, time-lapse  
20 profiles, which were obtained during the induction of differentiation under the above-described conditions, were compared with each other. The time-lapse profiles were obtained as follows. Each reporter gene was introduced into mesenchymal stem cells by a solid phase transfection method.  
25 The cells were cultured in undifferentiated state maintenance medium for two days. Thereafter, the medium was replaced with osteoblast differentiation medium. This time point was referred to as the osteoblast differentiation start time. Supplement factors were added at  
30 concentrations recommended for the osteoblast

differentiation medium. The other culture conditions were in accordance with Cambrex's instructions.

The results are shown in Figure 26. The profile pattern on the left of Figure 26 was obtained 10 hours to 30 hours after replacement of the medium. The profile pattern on the right of Figure 26 was obtained 5 to 6 days after replacement of the medium. Thus, it was demonstrated that the pattern significantly changed over time. The phases of the profiles were calculated using a formula shown in Figure 27 and the results were summarized in a table to the right of Figure 27. As can be seen, the inversion of the phase of the profile was deeply associated with differentiation for ISRE, RARE, STAT3, GRE, CRE, TRE, E2F, and p53. Therefore, it was demonstrated that by examining the phase, changes in process, i.e., the occurrence of transcription control, could be detected.

(Arbitrary combination of reporters)

Next, it was demonstrated that differentiation could be identified using an arbitrary combination of promoters for which data was extracted at the initial stage of induction of differentiation. Briefly, the analysis was conducted as shown in Figure 20.

The results are shown in Figure 20. This analysis revealed that although differentiation could not be detected at its very initial stage (potentially due to noise), but could be confirmed about 15 hours after induction of differentiation. In this example, when data

was extracted for 8 or more promoters, differentiation could be detected at a detection rate of 100%. When data was extracted for 3 promoters, differentiation could be detected at a detection rate of more than 90%. When data was extracted for two promoters, differentiation could be detected at a detection rate of 88%. When data was extracted for one promoter, differentiation could be detected at a detection rate of 82%. Thus, it was revealed that one, two or at least three promoters are sufficient for determination or identification of the state of cells.

(Maintenance of undifferentiated state)

Next, the maintenance of undifferentiated state was analyzed using an arbitrary combination of transcription control sequences for which data was extracted. Analysis was conducted as described in Figure 20.

The results are shown in Figure 21. As is largely different from the results of induction of differentiation, by comparing the profiles of the transcription control sequences with one another, it could be determined whether or not stem cells were induced into differentiation or remained undifferentiated. Such a determination could be achieved using at least one transcription control sequence. The determination of the state of cells using such a small number of transcription control sequences cannot be achieved by conventional techniques. It can be said that the present invention achieved an excellent effect.

30

By analyzing a cellular process in such a fashion, the formation of cellular functions can be described as a cocktail party process as shown in Figure 22. With such a process description, the present invention made  
5 it possible to analyze the progression of a cellular response to drugs and the progression of the induction of differentiation.

(Example 5 Anticancer agent)

10 In this example, cisplatin was used as an exemplary anticancer agent and was mixed into the medium of exposed cells. The concentration of the anticancer agent was selected as appropriate, such as 1  $\mu$ M, 5  $\mu$ M, 10  $\mu$ M, and the like, to observe the reaction of the cells. Cisplatin  
15 was applied to cells resistant or sensitive to the anticancer agent. Time-lapse observation was conducted to produce profiles as in the above-described examples. As a result, it was revealed that time-lapse profiles varied depending on the difference in cisplatin concentration and  
20 resistance/sensitivity.

(Example 6: RNAi)

The present Example demonstrated that it was possible to obtain a profile relating to gene knockdown  
25 effect using a cell was immobilized as described in Example 1, RNAi was used as a biological agent. The following was used as RNAi for experimentation. Gene expression inhibition methods using ribozymes and siRNA and the like allows obtaining profiles of response reactions in a cell  
30 for which gene expression inhibition is conducted using the

same.

RNAi: those sequences available at the URL:  
<http://www.nippongene.jp/pages/products/sirna/review/>  
5 were used (for example, Control siRNA duplex).

(RNAi transfection)

First, it was confirmed whether the siRNA could  
achieve knockdown effects. Synthesis of siRNA  
10 5'-AAGCAGCAGGACUUCUUAAG-3' (SEQ ID NO: 2) corresponding  
to EGFP was performed to prepare an array substrate as  
described herein above in the Examples. The preparation of  
array substrate using siRNA instead of nucleic acid  
molecules including promoter sequences was performed.  
15 Transfection using these array substrates confirmed  
effective inhibition of expression of a target gene. The  
protocols thereof are presented in Figure 28.

(Results)

20 Figure 29A shows the effects of target gene inhibition  
by siRNA. Expression of the target gene has actually been  
inhibited. The results using this gel may be stored as a  
profile in any data format.

25 Next, results of siRNA are stored as a profile data  
(image data of TIFF format having resolution at the level  
of 5  $\mu$ m/pixel or less). As such, the results of siRNA may  
be stored as a profile data. Such a format is not limited  
to those specifically presented in this Example, but those  
30 skilled in the art may employ any type of formats.

(Figure 9: Applications using siRNA and a transfection microarray of PC12 cells on a collagen IV coated chip)

5               Next, the present Example conducted a gene expression inhibition experiment using siRNA. The present Example evaluated whether or not the present invention functions by observing whether or not siRNAs against EGFP can specifically inhibit the expression of the EGFP as an  
10   indicator.

              Using the conditions described in Example 7, transfection of PC12 was conducted on an array coated with collagen IV. In lieu of the gene used in Example 7, the  
15   following conditions were used:

              0.75 ng of an expression vector (pEGFP-N1), HcRed (available from BD Clontech) were each spotted on a single, specific spot of the array. Thereafter, 16.5 ng of siRNA  
20   (available from Dharmacon, target sequence: 5'-GGC TAC GTC CAG GAG CGC ACC -3' (SEQ ID NO:47)=a) or scrambled siRNA(available from Dharmacon. target sequence: 5'-gCg CgC TTT gTA ggA TTC g-3' (SEQ ID NO: 48)=b) were also spotted.

25               Figure **29B** shows the results. As shown in Figure **29 B(A)**, in the case of PC12 cells co-transfected with EGFP vector and anti-EGFP siRNA, it was observed that only signals from HcRed were detected, and the green signal which should be derived from pEGFP-N1 had been inhibited.  
30   On the other hand, as shown in Figure **29 B(B)**, in the case

of scrambled siRNA, green fluorescence was been observed and thus it was confirmed that the effects seen in Figure 29 B(A) is the result of RNAi. Relative intensities of the fluorescence in Figures 29B(A) and 29B(B) are shown in Figure 29B(C). y-axis is shown with relative luminance. It can be seen that the effect by EGFP is substantially completely inhibited.

Figure 29C shows a result and graph summarizing the above. The left-hand panel shows a photograph comparing an EGFP RNAi and a scrambled (Mock) RNAi. As shown in the figure, the use of RNAi of EGFP showed an inhibitory effect, whereas the use of scrambled RNAi did not show such an inhibitory effect. The right-hand panel shows the same together with DsRed2. Experimental conditions are similar to the above Examples. As a result, red (signal derived from DsRed) and green (signal derived from EGFP) were presented in the proportion of the effects of RNAi.

Figure 29D shows an illustrative drawing of a chip using an RNAi reporter. When using RNAi as an input signal, and introducing a nucleic acid encoding both a gene product capable of signaling such as EGF and the like, and a gene of interest (including a promoter) as an output, observation of the signaling as the output allows one to produce cellular information.

Figure 29E shows an exemplary experiments using a variety of reporters (pAP1-EGFP, pAP1(PMA)-EGFP, pCRE-EGFP, pE2F-EGFP, pERE-EGFP, pGAS-EGFP, pGRE-EGFP,



pHSE-EGFP, pISRE-EGFP, pMyc-EGFP, pNFAT-EGFP, pNFkB-EGFP, pRARE-EGFP, pRb-EGFP, pSTST3-EGFP, pSRE-EGFP, pTRE-EGFP, pp53-EGFP, pCREB-sensor, pIkB-sensor, pp53-sensor, pCasapase3-sensor; cis-element sequence was commercially  
5 available from Clontech; these are plasmid vectors produced by recombining a fluorescent protein gene) . As such, the system of the present invention will function regardless of the types of reporters used.

10 (Example 7: Regulation of gene expression using a tetracycline-dependent promoter)

As described in the Examples 1-3, it was demonstrated that a tetracycline-dependent promoter could be used to produce a profile showing how gene expression  
15 is regulated. The sequences described below were used.

As the tetracycline-dependent promoter (and its gene vector construct), pTet-Off and pTet-On vectors (BD Biosciences) were used (see  
20 <http://www.clontech.com/techinfo/vectors/cattet.shtml>). As a vector, pTRE-d2EGFP (SEQ ID NO.: 29) was used (see <http://www.clontech.com/techinfo/vectors/vectorsT-Z/pTRE-d2EGFP.shtml>).

25 (Protocol)

pTet-Off and pTet-On (SEQ ID NOS.: 26 and 27, respectively) were printed onto array substrates. Real time measurement was performed on the array substrates to determine whether or not tetracycline regulates gene  
30 expression. The results are shown in Figure 30. As shown

in Figure 30, a change in gene expression was detected only for the tetracycline-dependent promoter. Figure 31 is a photograph showing the actual states of expression for the tetracycline-dependent promoter and the  
5 tetracycline-independent promoter. As can be seen, the difference between them is measurable by the naked eye.

(Measurement of profile data)

Images are taken in real time. Changes in  
10 intensity per cell or area are plotted on a graph. The resultant data may be subjected to linear transformation, such as noise reduction, and then multivariate analysis, signal processing, or the like, to obtain profile data. The resultant data is compared between phenomena or cells,  
15 thereby making it possible to determine a specific response or identity for the cells.

(Example 8: Gene expression)

Next, nucleic acid molecules encoding  
20 structural genes were used to produce cellular profiles. In this example, an olfactory receptor I7 (SEQ ID NOS: 13, 14) was used as a structural gene. The protocol used in Examples 1-3 was used.

25 As a result, as with promoters, it was demonstrated that cellular profiles could be produced by measuring the amount of gene products or the like.

(Example 9: Apoptotic signals)

30 Next, it was investigated that cellular

profiles could be produced by monitoring the activation of caspase 3 present within cells. Transfection and array preparation were performed as in the above-described examples.

5

pCaspase3-Sensor Vector (BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303; cat. No. 8185-1) was used to monitor an apoptotic signal from caspase 3.

10

As a result, as with promoters, it was demonstrated that cellular profiles could be produced by measuring apoptotic signals or the like.

15

(Example 10: Stress signal)

Next, it was investigated whether cellular profiles characterizing stress signals from JNK, ERK, p38 or the like could be produced using transcription factor reporters. Transfection and array preparation were performed as in the above-described examples.

20

pAP1-EGFP, pCRE-EGFP, and pSRE-EGFP available from BD Bioscience Clontech were used to monitor stress signals from JNK, ERK, and p38.

25

As a result, as in the above-described examples, it was demonstrated that cellular profiles could be produced by measuring stress signals.

30

(Example 11: Localization of molecules)

Next, it was demonstrated that a gene of interest could be fused with a fluorescent protein so that the expression profile of the gene product and the cellular localization of the gene product could be visualized.

5

GFP, RFP, CFP and BFP, were used as fluorescent proteins and cloned KIAA cDNA libraries or the like were used as genes of interest to produce gene constructs. These materials are specifically described below:

10

cloned KIAA cDNA (KIAA=Kazusa DNA Research Institute, Kazusa, Chiba, Japan); and  
cDNA libraries commercially available from Invitrogen.

15

Transfection and array preparation were performed as in the above-described examples.

The expression of cloned KIAA, KIAA1474, was monitored to produce a profile of the expression and investigate the localization of the expression product.

20

As a result, as in the above-described examples, it was demonstrated that intentionally constructed gene constructs could be used to produce cellular profiles for target characteristics.

25

(Example 12: Changes in cellular morphology)

Next, it was demonstrated that cellular profiles characterizing cellular morphology could be

30

produced by expressing or knocking out genes or adding substances (glycerophosphate as a chemical substance and dexamethasone as a cytokine). Cellular morphology, such as multinucleated cells, cellular outgrowth, outgrowth  
5 projections, and the like, was measured and analyzed as three-dimensional data.

The specific sequences of the introduced nucleic acid molecules are described below:

10

Cloned KIAA (*supra*); and  
RNAi for transcription factors (CBFA-1, AP1).

Transfection and array preparation were  
15 performed as in the above-described examples.

Mesenchymal stem cells as used in the above-described examples were used to monitor the morphology of cells which were induced to be differentiated into  
20 osteoblasts.

As a result, as in the above-described examples, it was demonstrated that intentionally constructed gene constructs could be used to produce cellular profiles for  
25 target characteristics. Event descriptors can be produced based on the profile data using the process as used in the above-described examples.

(Example 13: Intermolecular interaction)

30

Next, it was demonstrated that cellular

profiles could be produced by using a technique such as a two-hybrid system, FRET, BRET, or the like.

5 The specific sequences of the introduced nucleic acid molecules are described below:

olfactory receptors (SEQ ID NOS: 13 to 38); and G proteins (SEQ ID NOS: 39 to 44).

10 Transfection and array preparation were performed as in the above-described examples.

15 The dissociation of the olfactory receptor and G protein was monitored through induction of a scented substance, which was captured as changes in fluorescent wavelength. In this manner, cells were monitored.

The two-hybrid system, FRET, and BRET were specifically performed as follows.

20

The two-hybrid system was available from Clontech

(<http://www.clontech.co.jp/product/catalog/007003006.shtml>). FRET and BRET were performed using devices available

25 from Berthold Japan.

30 As a result, as in the above-described examples, it was demonstrated that intentionally constructed gene constructs could be used in a two-hybrid system, FRET, BRET, or the like, to produce cellular profiles.

## (EXAMPLE 14: Receptor-Ligand)

Next, it was demonstrated that a cellular profile can be produced by employing the interaction between a receptor and its ligand as an indicator. It is useful for network formation in a cell, to obtain interactive information between a receptor protein present in the cell membrane or nuclear membrane, or the like, and a ligand thereto.

10

In the present Example, the following was prepared:

## (Cell adhesion molecules)

15

A variety of extracellular matrix protein and variants and fragments thereof were prepared as candidates for cell adhesion molecules. What was prepared in the present Example is as follows. Cell adhesion molecules were commercially available.

20

- 1) ProNectin F (Sanyo Chemical Industries, Kyoto, Japan);
- 2) ProNectin L (Sanyo Chemical Industries);
- 3) ProNectin Plus (Sanyo Chemical Industries);
- 4) fibronectin (SEQ ID NO.: 2);
- 25 5) gelatin.

Plasmids were prepared as DNA for transfection. Plasmids, pEGFP-N1 and pDsRed2-N1 (both from BD Biosciences, Clontech, CA, USA) were used. In these plasmids, gene expression was under the control of cytomegalovirus (CMV).

30

The plasmid DNA was amplified in E. coli (XL1 blue, Stratgene, TX, USA) and the amplified plasmid DNA was used as a complex partner. The DNA was dissolved in distilled water free from DNase and RNase.

5

The following transfection reagents were used: Effectene Transfection Reagent (cat. no. 301425, Qiagen, CA), TransFast<sup>TM</sup> Transfection Reagent (E2431, Promega, WI), Tfx<sup>TM</sup>-20 Reagent (E2391, Promega, WI), SuperFect Transfection Reagent (301305, Qiagen, CA), PolyFect Transfection Reagent (301105, Qiagen, CA), LipofectAMINE 2000 Reagent (11668-019, Invitrogen corporation, CA), JetPEI (×4) conc. (101-30, Polyplus-transfection, France), and ExGen 500 (R0511, Fermentas Inc., MD). These  
10 transfection reagents were added to the above-described DNA and actin-like acting substance in advance or complexes thereof with the DNA were produced in advance.

The thus-obtained solution was used in assays  
20 using transfection arrays described below. Next, transfection effects on a solid phage were observed. The protocols therefor are described below:

(Protocol)

25 The final concentration of DNA was adjusted to 1 µg/µL. A cell adhesion molecule was preserved as a stock having a concentration of 10 µg/µL in ddH<sub>2</sub>O. All dilutions were made using PBS, ddH<sub>2</sub>O, or Dulbecco's MEM. A series of dilutions, for example, 0.2 µg/µL, 0.27 µg/µL, 0.4 µg/µL,  
30 0.53 µg/µL, 0.6 µg/µL, 0.8 µg/µL, 1.0 µg/µL, 1.07 µg/µL,



1.33  $\mu\text{g}/\mu\text{L}$ , and the like, were formulated.

Transfection reagents were used in accordance with instructions provided by each manufacturer.

5

Plasmid DNA was removed from a glycerol stock and amplified in 100 mL L-amp overnight. Qiaprep Miniprep or Qiagen Plasmid Purification Maxi was used to purify DNA in accordance with a standard protocol provided by the  
10 manufacturer.

In the present Example, the following five cells were used to confirm an effect: human mesenchymal stem cell (hMSCs, PT-2501, Cambrex BioScience Walkersville, Inc.,  
15 MD); human embryonic renal cell (HEK293, RCB1637, RIKEN Cell Bank, JPN); NIH3T3-3 cell (RCB0150, RIKEN Cell Bank, JPN); HeLa cell (RCB0007, RIKEN Cell Bank, JPN); and HepG2 (RCB1648, RIKEN Cell Bank, JPN). These cells were cultured in DMEM/10% IFS containing L-glut and pen/strep.

20

(Dilution and DNA spots)

Transfection reagents and DNA were mixed to form a DNA-transfection reagent complex. The complex formation requires a certain period of time. Therefore, the mixture  
25 was spotted onto a solid phase support (e.g., a poly-L-lysine slide) using an arrayer. In the present Example, as a solid phase support, an APS slide, a MAS slide, and an uncoated slide were used, as well as a poly-L-lysine slide. These slides are available from Matsunami Glass (Kishiwada, Japan)  
30 or the like.

For complex formation and spot fixation, the slides were dried overnight in a vacuum dryer. Drying was performed for a duration in the range of 2 hours to 1 week.

5

Although the cell adhesion molecule might be used during the complex formation, it was also used immediately before spotting in the present Example.

10 (Formulation of mixed solution and application to solid phase supports)

300  $\mu$ L of DNA concentrated buffer (EC buffer) + 16  $\mu$ L of an enhancer were mixed in an Eppendorf tube. The mixture was mixed with a Vortex, followed by incubation for 15 5 minutes. 50  $\mu$ L of a transfection reagent (Effectene, etc.) was added to the mixture, followed by mixing by pipetting. To apply a transfection reagent, an annular wax barrier was formed around the spots on the slide. 366  $\mu$ L of the mixture was added to the spot region surrounded by 20 the wax, followed by incubation at room temperature for 10 to 20 minutes. Thereby, the fixation to the support was manually achieved.

(Distribution of cells)

25 Next, a protocol for adding cells will be described. Cells were distributed for transfection. The distribution was typically performed by reduced-pressure suction in a hood. A slide was placed on a dish, and a solution containing cells was added to the dish for 30 transfection. The cells were distributed as follows.

The growing cells were seeded at a concentration of  $10^7$  cells/25 mL. The cells were plated on the slide in a 100×100×15 mm squared Petri dish or a 100 mm (radius) × 15 mm circular dish. Transfection was conducted for about 40 hours. This period of time corresponded to about 2 cell cycles. The slide was treated for immunofluorescence.

(Evaluation of gene introduction)

Gene introduction was evaluated by detection using, for example, immunofluorescence, fluorescence microscope examination, laser scanning, radioactive labels, and sensitive films, or emulsion.

When an expressed protein to be visualized is a fluorescent protein, such a protein can be observed with a fluorescence microscope and a photograph thereof can be taken. For large-sized expression arrays, slides may be scanned using a laser scanner for storage of data. If an expressed protein can be detected using specific fluorescence in the case of calcium, a protocol specific for detection of a specific fluorescence can be successively performed to detect signals. If an expressed protein can be detected using fluorescence antibodies, an immunofluorescence protocol can be successively performed.

(Laser scanning and Quantification of fluorescence intensity)

To quantify transfection efficiency, the present inventors used a DNA microarray scanner (GeneTAC

UC4x4, Genomic Solutions Inc., MI). Total fluorescence intensity (arbitrary units) was measured, and thereafter, fluorescence intensity per unit surface area was calculated.

5 (Cross-sectional observation by confocal scanning microscope)

Cells were seeded on tissue culture dishes at a final concentration of  $1 \times 10^5$  cells/well and cultured in appropriate medium (Human Mesenchymal Cell Basal Medium  
10 (MSCGM BulletKit PT-3001, Cambrex BioScience Walkersville, Inc., MD). After fixation of the cell layer with 4% paraformaldehyde solution, SYTO and Texas Red-X phalloidin (Molecular Probes Inc., OR, USA) was added to the cell layer for observation of nuclei and F-actin. The samples emitting  
15 light due to gene products and the stained samples were observed with a confocal laser microscope (LSM510: Carl Zeiss Co., Ltd., pinhole size=Ch1=123  $\mu\text{m}$ , Ch2=108  $\mu\text{m}$ , image interval = 0.4) to obtain cross sectional views.

20 Next, an Example, to which the present invention is applied to, is described wherein an olfactory receptor is selected as a typical example of a chemical substance receptor. When a preliminary example was implemented, it was proved that transfection arrays can also be used for  
25 an olfactory receptor

The olfactory receptor expression vector group was spotted per every kind of receptor, on a cover glass, which was made like an array, was secured with screws and  
30 the like in a chamber for signal measurement, and cells

having an almost homogeneous nature, were cultured thereon. Regarding a chamber for signal measurement, sample gas was introduced into a known structure (Proc. Natl. Acad. Sci. USA, 96(1999): 4040-4045 and the like). Other devised  
5 chambers are also intended. During response measurement, culture medium was passed through the chamber at a constant speed. Culture media was supplied to the chamber for measurement from the opening of a culture medium supply tube, and a sample gas supplying tube was secured at the position  
10 preferably near the liquid level, which is the upper portion of an interval whose boundary is defined by reaching a wall which prevents the approach of culture media over a cover-slip that forms the ceiling of the measurement member, so that sample gas can be supplied to culture medium flowing  
15 across the interval. This sample gas supplying tube is preferably made of materials to which lipophilic odor substances such as Teflon and peak, and dust are not readily adsorbed. The higher effect was obtained in the situation wherein, at the time other than introducing sample gas,  
20 sample gas remaining in a tube was removed, and to preferably keep the interior clean, the tube (preferably with a broad opening) could be purged with odorless air by setting a three-way valve in the mid course, or by setting a check valve at a joint of an odorless air supplying tube. However,  
25 it was not necessary. The example could also be implemented in the situation wherein, at a time other than when introducing sample gas from outside for an appropriate time such as 0.5-4 seconds, odorless air was introduced from mid course of a sample gas supplying tube near a opening for  
30 collecting gas from outside, the interior of the tube was

washed therewith, and at the same time, odorless gas was supplied to the culture medium as sample gas to promote the removal of remaining gas in a measurement chamber. A supporting base for the upper-glass cover slip is made of  
5 a water repellent opaque plastic such as Teflon. A width of flow channel, where culture medium flows, is about 2-fold of a width of an array, and the array is disposed in the center of the flow channel. Regarding a culture medium supplying tube and an overflow culture medium sucking tube,  
10 a portion several millimeters from the opening at the side of the measurement chamber is made using materials which have high hydrophilicity and are difficult to deform, such as stainless steel. The upper portion of the supporting base of the upper glass cover-slip where culture medium flows,  
15 from the openings of both tubes to an array, was coated, or covered with a pieces of lens paper and the like in order to provide sufficient hydrophilicity. Negative pressure for suction was adjusted at the grade such that measurements were not affected by vibration from sound generated by  
20 aspiration culture.

Generally, response measurement could be implemented 2 days after the gene introduced by the vector was expressed. Since an upper glass cover-slip was required  
25 only at the time of measurement, it was not required to install it during culture until the gene was expressed. Therefore, the Example could be implemented, adding an upper glass cover slip integrated with a wall which prevents leakage of culture medium, and a supporting base for the  
30 upper glass cover slip, to a chamber for measurement, when

setting a chamber for measurement of a change in fluorescence measured by an apparatus after the gene expressed. The Example could also be implemented in the situation wherein culture medium was exchanged without using a culture medium supply tube and an overflow culture suction line tube during culture until the gene was expressed. An amount of about 10ml of culture medium was supplied and exchanged at the frequency of about once per several hours to one day, only during the time tissue culture was performed.

10

The size of odor response could be optically measured using a two-dimensional image sensor such as a sensitive video camera, with a calcium ion sensitive fluorescent dye fura-2 and the like absorbed into the cell. The measurement interval preferably has time resolution which can evaluate time constants of build-up and recovery of response of about 1/3-1 second. However, if average response time curve or its theoretical formula had been obtained, actual change was estimated from measurement results at 5 points with 5-second-intervals, 5, 10, 15, 20, and 25 seconds after stimulation. The obtained estimates of the time constant of the response starting time, response build-up time, and response recovery time was set as an index, and evaluation could be made as to whether a signal was induced by odor, or generated by spontaneous activity of a cell, or other abnormalities.

In this Example, the response of an expressed olfactory receptor in olfactory receptor neuron was studied by measuring the change in fluorescence intensity of calcium

30

sensitive fluorescent dye. Decrease of fluorescence intensity (downward change) corresponds to the response of an olfactory receptor. Odor molecules were added to the culture at the concentration indicated above them as stimulation source, and administered to a cell during the time indicated by a bar (4 or 2 seconds). As understood from this example, responses measured simultaneously in a simultaneously adjusted cell have high intercommunity in response time characteristics, response threshold concentration corresponding to different stimulation per cell, and relative value of response amplitude. However, cells adjusted at a different times show some differences. These results show that the highest measurement reliability can be obtained by measuring odor response using a sensor arrayed to a size that allows a homogeneous administration of sample gas, providing the same adjustment conditions.

(Example 15: Application to neuron differentiation)

Next, experiments similar to those of Example 14 have been conducted with neurons to analyze the effects of tyrosine kinase RNAi with a transfection microarray. The exemplified drawings are shown in Figure **31B**.

As shown in Figure **31B**, network analysis can be conducted by taking photographs of signal represented by a reporter and collecting information thereon.

Figure **31C** shows the responses of retinoic acid (RA) and nerve growth factor (NGF) to a variety of tyrosine



kinases. Inhibition % by siRNA is shown.

Figure **31D** depicts an exemplary drawing of a signal transduction pathway obtained as a result of analysis.

5

Figure **31E** shows results obtained by the above-mentioned analysis. Classification has been made regardless of whether the subject cells were dopaminergic neurons, cholinergic neuron, both, or neither. It can be  
10 concluded that those kinases relating to both cell types have high probability of relating to nerve projection formation.

(Example 16: Data production)

15

Data produced in Examples 5-15 can be analyzed using a mathematical analysis with an appropriate modification as described in Example 4. Such data have been presented in a variety of formats.

20

(Example 17: Production of a digital cell)

Data produced in Examples 5-15 and additional data produced using the protocols described therein were used to produce a digital cell. In order to produce digital cells, parameters for data produced in these Examples have  
25 been extracted, and medium, pH, temperature, CO<sub>2</sub> concentration, and the like have been used as environment parameters. Database production may be performed using, for example, a spreadsheet software such as Excel™ available from Microsoft, or a database software such as Access™ also  
30 available from Microsoft. Next, as cell parameters, a

database including cell species such as those used in Examples 5-15 can be used. A variety of stimulus parameters such as a variety of chemical stimuli (for example, including a variety of growth factors or cytokines such as HGF, FGF, PDGF, VEGF, CSF and the like) can be input to produce cell dynamics data, measurement data of reporters such as fluorescence intensity and the like. As such, a database constituting digital cell can be produced. Such examples are shown in Figures **33A** and **33B**.

10

(Example 18: Use of digital cells: *in silico* live experiments)

The digital cells produced in Example 17 have been used to conduct experiments on a computer. In the present Example, a mesenchymal stem cell is used to study which agents are differentiation agents. In the case of Figure **33A**, cell A is selected as cell (for example, mesenchymal stem cell or the like). Further, DMEM is selected as a medium, pH 7.4 is selected as the pH, 37 degree Celcius is selected as the temperature, 5% is selected as the CO<sub>2</sub> concentration. Moreover, a variety of chemical stimuli such as growth factors or cytokines such as HGF, FGF, PDGF, VEGF, and CSF are selected. With respect to such a variety of chemical stimuli, concentrations are also appropriately selected, such as 1nM to 1mM. Combinations of these two or three thereof are also selected as a variety of chemical stimuli. Depending these combination and concentration, data regarding responses with respect to how a mesenchymal stem cell responds is output. As an output, cell dynamic is included. From such cell dynamic, it is

30

confirmed that the mesenchymal stem cell is differentiated (e.g. to bone marrow or adipocyte or the like) or not. If morphology is not sufficient, combinations between transcriptional factors and EGF as reporters are used to output further measurement data. As such, it can be confirmed whether or not a mesenchymal stem cell is differentiated to which a specific differentiated cell. Using the present method, one can specify a chemical stimulus which induces differentiation to a specific differentiated cell.

10

(Example 19: Use of digital cells - education by *in silico* live experiments

*In silico* live experiments described in Example 18 are conducted during school education. In this example, the experimental theme as described above is given to a student. The student selects a variety of parameters from a database of a given digital cells. The student composes his/her own research based on the data selected. The student submits the composed research results as assignment/report. As such, education to a student can be conducted without using a live experimental system.

15

(Example 20: provision of a digital service)

A database of the digital cell may be provided as an external service. Databases produced in Example 18 may use those embodiment described in Figure 35. As such, the configuration of computer system 3501 providing a service reproducing experimental results to an actual cell using the digital cells is shown. Computer system 3501 comprises service requester 3510 requesting services

20

25

30

desired by a user, and service provider 3520 providing a determined service in response to the request. Users such as research institutes, educational organizations or institutions request desired services. Service provider 5 3510 providing commercial service provides appropriate data to the research institutes, educational organizations or institutions upon request. For the purpose of school education, for example, a particular data base only directed to a particular cell or parameters or the like may be of 10 service target.

As such, it is demonstrated that the digital cell of the present invention can be used to provide services.

15 Although certain preferred embodiments have been described herein, it is not intended that such embodiments be construed as limitations on the scope of the invention except as set forth in the appended claims. 20 Various other modifications and equivalents will be apparent to and can be readily made by those skilled in the art, after reading the description herein, without departing from the scope and spirit of this invention. All patents, published patent applications and publications cited herein are 25 incorporated by reference as if set forth fully herein.

#### INDUSTRIAL APPLICABILITY

According to the present invention, it is 30 possible to determine the state of cells by observing a

surprisingly small number of factors. Therefore, the present invention is applicable to diagnosis, prevention, and treatment. The present invention is also applicable to the fields of food, cosmetics, agriculture, environmental  
5 engineering, and the like. As live experiments can be reproduced on a computer, education and research in the field of biotechnology can be conducted on such a computer, which is industrially applicable.